

# Optimal Uncertainty Quantification

H. Owhadi, T. J. Sullivan, M. McKerns, M. Ortiz  
California Institute of Technology  
owhadi@caltech.edu

C. Scovel  
Los Alamos National Laboratory  
jcs@lanl.gov

May 25, 2012

## Abstract

We propose a rigorous framework for Uncertainty Quantification (UQ) in which the UQ objectives and the assumptions/information set are brought to the forefront. This framework, which we call *Optimal Uncertainty Quantification* (OUQ), is based on the observation that, given a set of assumptions and information about the problem, there exist optimal bounds on uncertainties: these are obtained as values of well-defined optimization problems corresponding to extremizing probabilities of failure, or of deviations, subject to the constraints imposed by the scenarios compatible with the assumptions and information. In particular, this framework does not implicitly impose inappropriate assumptions, nor does it repudiate relevant information.

Although OUQ optimization problems are extremely large, we show that under general conditions they have finite-dimensional reductions. As an application, we develop *Optimal Concentration Inequalities* (OCI) of Hoeffding and McDiarmid type. Surprisingly, these results show that uncertainties in input parameters, which propagate to output uncertainties in the classical sensitivity analysis paradigm, may fail to do so if the transfer functions (or probability distributions) are imperfectly known. We show how, for hierarchical structures, this phenomenon may lead to the non-propagation of uncertainties or information across scales.

In addition, a general algorithmic framework is developed for OUQ and is tested on the Caltech surrogate model for hypervelocity impact and on the seismic safety assessment of truss structures, suggesting the feasibility of the framework for important complex systems.

The introduction of this paper provides both an overview of the paper and a self-contained mini-tutorial about basic concepts and issues of UQ.

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	The UQ problem . . . . .	4
1.2	Motivating physical example and outline of the paper . . . . .	6
1.2.1	Formulation of the admissible set and reduction theorems. . . . .	8
1.2.2	Optimal concentration inequalities. . . . .	10
1.2.3	Computational framework. . . . .	11
1.2.4	Outline of the paper. . . . .	15
<b>2</b>	<b>Optimal Uncertainty Quantification</b>	<b>15</b>
2.1	First description . . . . .	16
2.2	The Optimal UQ loop . . . . .	21
<b>3</b>	<b>Generalizations and Comparisons</b>	<b>29</b>
3.1	Prediction, extrapolation, verification and validation . . . . .	29
3.2	Comparisons with other UQ methods . . . . .	30
3.3	OUQ with random sample data . . . . .	33
<b>4</b>	<b>Reduction of OUQ Optimization Problems</b>	<b>34</b>
4.1	Reduction of OUQ . . . . .	35
4.2	Generalized moments of the response function . . . . .	38
4.3	Application to McDiarmid’s inequality . . . . .	40
<b>5</b>	<b>Optimal Concentration Inequalities</b>	<b>43</b>
5.1	Explicit solutions under the assumptions of McDiarmid’s inequality . . . . .	43
5.1.1	Explicit solutions in dimensions one and two . . . . .	44
5.1.2	Explicit solution in dimension three . . . . .	44
5.1.3	Solution in dimension $m$ . . . . .	47
5.2	Explicit solutions under the assumptions of Hoeffding’s inequality . . . . .	48
<b>6</b>	<b>Computational Implementation</b>	<b>49</b>
6.1	Extreme points of reduced OUQ problems are attractors . . . . .	49
6.2	Coagulation–Fragmentation algorithm for OUQ . . . . .	49
6.3	The OUQ algorithm in the <i>mystic</i> framework . . . . .	53
<b>7</b>	<b>Application to the Seismic Safety Assessment of Structures</b>	<b>54</b>
7.1	Formulation in the time domain . . . . .	55
7.1.1	Formulation of the problem . . . . .	55
7.1.2	Formulation of the information set . . . . .	56
7.1.3	The OUQ optimization problem . . . . .	58
7.1.4	Reduction of the optimization problem . . . . .	58
7.1.5	Numerical results . . . . .	58
7.2	OUQ and critical excitation. . . . .	62
7.3	Alternative formulation in the frequency domain . . . . .	63

7.3.1	Formulation of the information set . . . . .	63
7.3.2	OUQ objectives . . . . .	64
7.3.3	Reduction of the optimization problem with Dirac masses . . . . .	65
7.3.4	Reduction of the optimization problem based on strong duality . . . . .	65
7.3.5	Numerical results . . . . .	66
<b>8</b>	<b>Application to Transport in Porous Media</b>	<b>67</b>
<b>9</b>	<b>Conclusions</b>	<b>69</b>
	<b>Acknowledgements</b>	<b>71</b>
<b>10</b>	<b>Appendix: Proofs</b>	<b>71</b>
10.1	Proofs for Section 4 . . . . .	71
10.2	Proofs for Section 5 . . . . .	75
	<b>References</b>	<b>83</b>

# 1 Introduction

## 1.1 The UQ problem

This paper sets out a rigorous and unified framework for the statement and solution of uncertainty quantification (UQ) problems centered on the notion of available information. In general, UQ refers to any attempt to quantitatively understand the relationships among uncertain parameters and processes in physical processes, or in mathematical and computational models for them; such understanding may be deterministic or probabilistic in nature. However, to make the discussion specific, we start the description of the proposed framework as it applies to the certification problem; Section 3 gives a broader description of the purpose, motivation and applications of UQ in the proposed framework and a comparison with current methods.

By *certification* we mean the problem of showing that, with probability at least  $1 - \epsilon$ , the real-valued response function  $G$  of a given physical system will not exceed a given safety threshold  $a$ . That is, we wish to show that

$$\mathbb{P}[G(X) \geq a] \leq \epsilon. \quad (1.1)$$

In practice, the event  $[G(X) \geq a]$  may represent the crash of an aircraft, the failure of a weapons system, or the average surface temperature on the Earth being too high. The symbol  $\mathbb{P}$  denotes the probability measure associated with the randomness of (some of) the input variables  $X$  of  $G$  (commonly referred to as “aleatoric uncertainty”).

Specific examples of values of  $\epsilon$  used in practice are:  $10^{-9}$  in the aviation industry (for the maximum probability of a catastrophic event per flight hour, see [83, p.581] and [15]), 0 in the seismic design of nuclear power plants [28, 23] and 0.05 for the collapse of soil embankments in surface mining [36, p.358]. In structural engineering [31], the maximum permissible probability of failure (due to any cause) is  $10^{-4}K_s n_d/n_r$  (this is an example of  $\epsilon$ ) where  $n_d$  is the design life (in years),  $n_r$  is the number of people at risk in the event of failure and  $K_s$  is given by the following values (with 1/year units): 0.005 for places of public safety (including dams); 0.05 for domestic, office or trade and industry structures; 0.5 for bridges; and 5 for towers, masts and offshore structures. In US environmental legislation, the maximum acceptable increased lifetime chance of developing cancer due to lifetime exposure to a substance is  $10^{-6}$  [53] ([44] draws attention to the fact that “there is no sound scientific, social, economic, or other basis for the selection of the threshold  $10^{-6}$  as a cleanup goal for hazardous waste sites”).

One of the most challenging aspects of UQ lies in the fact that in practical applications, the measure  $\mathbb{P}$  and the response function  $G$  are not known a priori. This lack of information, commonly referred to as “epistemic uncertainty”, can be described precisely by introducing  $\mathcal{A}$ , the set of all *admissible scenarios*  $(f, \mu)$  for the unknown — or partially known — reality  $(G, \mathbb{P})$ . More precisely, in those applications, the available information does not determine  $(G, \mathbb{P})$  *uniquely* but instead determines a set  $\mathcal{A}$  such that any  $(f, \mu) \in \mathcal{A}$  could a priori be  $(G, \mathbb{P})$ . Hence,  $\mathcal{A}$  is a (possibly infinite-dimensional) set of measures and functions defining explicitly *information on and assumptions about*  $G$

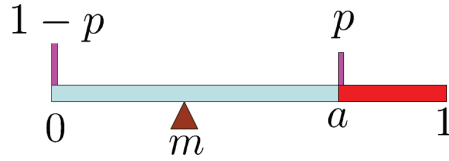


Figure 1.1: You are given one pound of play-dough and a seesaw balanced around  $m$ . How much mass can you put on right hand side of  $a$  while keeping the seesaw balanced around  $m$ ? The solution of this optimization problem can be achieved by placing any mass on the right hand side of  $a$ , exactly at  $a$  (to place mass on  $[a, 1]$  with minimum leverage towards the right hand side of the seesaw) and any mass on the left hand side of  $a$ , exactly at  $0$  (for maximum leverage towards the left hand side of the seesaw).

and  $\mathbb{P}$ . In practice, this set is obtained from physical laws, experimental data and expert judgment. It then follows from  $(G, \mathbb{P}) \in \mathcal{A}$  that

$$\inf_{(f, \mu) \in \mathcal{A}} \mu[f(X) \geq a] \leq \mathbb{P}[G(X) \geq a] \leq \sup_{(f, \mu) \in \mathcal{A}} \mu[f(X) \geq a]. \quad (1.2)$$

Moreover, it is elementary to observe that

- The quantities on the right-hand and left-hand of (1.2) are extreme values of optimization problems and elements of  $[0, 1]$ .
- Both the right-hand and left-hand inequalities are optimal in the sense that they are the sharpest bounds for  $\mathbb{P}[G(X) \geq a]$  that are consistent with the information and assumptions  $\mathcal{A}$ .

More importantly, in Proposition 2.1, we show that these two inequalities provide sufficient information to produce an optimal solution to the certification problem.

**Example 1.1.** To give a very simple example of the effect of information and optimal bounds over a class  $\mathcal{A}$ , consider the certification problem (1.1) when  $Y := G(X)$  is a real-valued random variable taking values in the interval  $[0, 1]$  and  $a \in (0, 1)$ ; to further simplify the exposition, we consider only the upper bound problem, suppress dependence upon  $G$  and  $X$  and focus solely on the question of which probability measures  $\nu$  on  $\mathbb{R}$  are admissible scenarios for the probability distribution of  $Y$ . So far, *any* probability measure on  $[0, 1]$  is admissible:

$$\mathcal{A} = \{\nu \mid \nu \text{ is a probability measure on } [0, 1]\}.$$

and so the optimal upper bound in (1.2) is simply

$$\mathbb{P}[Y \geq a] \leq \sup_{\nu \in \mathcal{A}} \nu[Y \geq a] = 1.$$

Now suppose that we are given an additional piece of information: the expected value of  $Y$  equals  $m \in (0, a)$ . These are, in fact, the assumptions corresponding to an elementary

Markov inequality, and the corresponding admissible set is

$$\mathcal{A}_{\text{Mrkv}} = \left\{ \nu \left| \begin{array}{l} \nu \text{ is a probability measure on } [0, 1], \\ \mathbb{E}_\nu[Y] = m \end{array} \right. \right\}.$$

The least upper bound on  $\mathbb{P}[Y \geq a]$  corresponding to the admissible set  $\mathcal{A}_{\text{Mrkv}}$  is the solution of the infinite dimensional optimization problem

$$\sup_{\nu \in \mathcal{A}_{\text{Mrkv}}} \nu[Y \geq a] \quad (1.3)$$

Formulating (1.3) as a mechanical optimization problem (see Figure 1.1), it is easy to observe that the extremum of (1.3) can be achieved only considering the situation where  $\nu$  is the weighted sum of mass a Dirac at 0 (with weight  $1 - p$ ) and a mass of Dirac at  $a$  (with weight  $p$ ). It follows that (1.3) can be reduced to the simple (one-dimensional) optimization problem: *Maximize  $p$  subject to  $ap = m$* . It follows that Markov's inequality is the optimal bound for the admissible set  $\mathcal{A}_{\text{Mrkv}}$ .

$$\mathbb{P}[Y \geq a] \leq \sup_{\nu \in \mathcal{A}_{\text{Mrkv}}} \nu[Y \geq a] = \frac{m}{a}. \quad (1.4)$$

In some sense, the OUQ framework that we present in this paper is the the extension of this procedure to situations in which the admissible class  $\mathcal{A}$  is complicated enough that a closed-form inequality such as Markov's inequality is unavailable, but optimal bounds can nevertheless be *computed* using reduction properties analogous to the one illustrated in Figure 1.1.

## 1.2 Motivating physical example and outline of the paper

Section 2 gives a formal description of the *Optimal Uncertainty Quantification* framework. In order to help intuition, we will illustrate and motivate our abstract definitions and results with a practical example: an analytical surrogate model for hypervelocity impact.

The physical system of interest is one in which a 440C steel ball (440C is a standard, i.e. a grade of steel) of diameter  $D_p = 1.778$  mm impacts a 440C steel plate of thickness  $h$  (expressed in mm) at speed  $v$  (expressed in  $\text{km} \cdot \text{s}^{-1}$ ) at obliquity  $\theta$  from the plate normal. The physical experiments are performed at the California Institute of Technology SPHIR (Small Particle Hypervelocity Impact Range) facility (see Figure 1.2). An analytical surrogate model was developed to approximate the perforation area (in  $\text{mm}^2$ ) caused by this impact scenario. The surrogate response function is as follows:

$$H(h, \theta, v) = K \left( \frac{h}{D_p} \right)^p (\cos \theta)^u \left( \tanh \left( \frac{v}{v_{\text{bl}}} - 1 \right) \right)_+^m, \quad (1.5)$$

where the *ballistic limit velocity* (the speed below which no perforation area occurs) is given by

$$v_{\text{bl}} := H_0 \left( \frac{h}{(\cos \theta)^n} \right)^s. \quad (1.6)$$

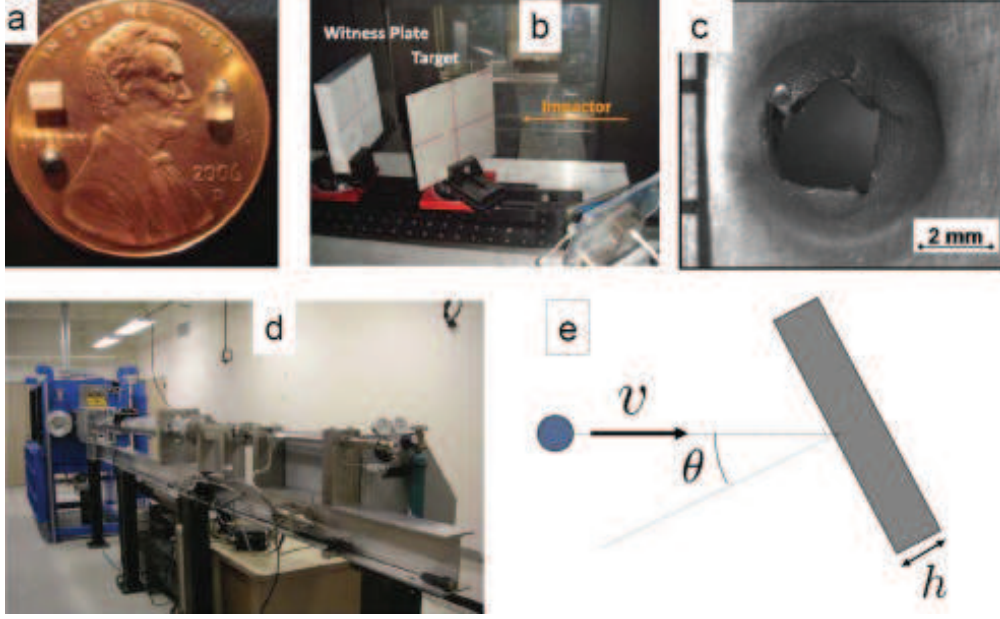


Figure 1.2: Experimental set up. (a) Stainless steel spherical projectiles and nylon sabots. (b) Target plate held at the end of the gun barrel (c) Perforation of the target plate (d) General view of the Small Particle Hypervelocity Impact Range (SPHIR) Facility at Caltech (e) plate thickness  $h$ , plate obliquity  $\theta$  and projectile velocity  $v$ .

The seven quantities  $H_0$ ,  $s$ ,  $n$ ,  $K$ ,  $p$ ,  $u$  and  $m$  are fitting parameters that have been chosen to minimize the least-squares error between the surrogate and a set of 56 experimental data points; they take the values

$$\begin{aligned} H_0 &= 0.5794 \text{ km} \cdot \text{s}^{-1}, & s &= 1.4004, & n &= 0.4482, & K &= 10.3936 \text{ mm}^2, \\ p &= 0.4757, & u &= 1.0275, & m &= 0.4682. \end{aligned}$$

Hence, in this illustrative example,  $H(h, \theta, v)$  will be our response function  $G(X_1, X_2, X_3)$  and we will consider cases in which  $H$  is perfectly and imperfectly known. In Section 7, we will apply the OUQ framework to the seismic safety assessment of structures and consider a more complex example involving a large number of variables.

### 1.2.1 Formulation of the admissible set and reduction theorems.

For the example considered here, we will assume that the input parameters  $h, \theta$  and  $v$  are random variables, of unknown probability distribution  $\mathbb{P}$  and of given range

$$\begin{aligned} (h, \theta, v) &\in \mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3, \\ h &\in \mathcal{X}_1 := [1.524, 2.667] \text{ mm} = [60, 105] \text{ mils}, \\ \theta &\in \mathcal{X}_2 := [0, \frac{\pi}{6}], \\ v &\in \mathcal{X}_3 := [2.1, 2.8] \text{ km} \cdot \text{s}^{-1}. \end{aligned} \tag{1.7}$$

We will measure lengths in both mm and mils (recall that 1 mm = 39.4 mils).

We will adopt the “gunner’s perspective” that failure consists of not perforating the plate, and therefore seek to obtain an optimal bound on the probability of non-perforation, i.e.  $\mathbb{P}[H \leq 0]$ , with possibly incomplete information on  $\mathbb{P}$  and  $H$ .

Assuming  $H$  to be known, if the information on  $\mathbb{P}$  is limited to the knowledge that velocity, impact obliquity and plate thickness are independent random variables and that the mean perforation area lies in a prescribed range  $[m_1, m_2] := [5.5, 7.5] \text{ mm}^2$ , then this information describes the admissible set  $\mathcal{A}_H$ , where

$$\mathcal{A}_H := \left\{ (H, \mu) \left| \begin{array}{l} H \text{ given by (1.5),} \\ \mu = \mu_1 \otimes \mu_2 \otimes \mu_3, \\ m_1 = 5.5 \text{ mm}^2 \leq \mathbb{E}_\mu[H] \leq m_2 = 7.5 \text{ mm}^2 \end{array} \right. \right\}. \tag{1.8}$$

If the information on  $H$  is limited to values of  $\text{Osc}_i(H)$ , the component-wise oscillations (defined below, it is a least upper bound on how a change in variable  $i$  affects the response function), and if the information on  $\mathbb{P}$  is as above, then the corresponding admissible set is  $\mathcal{A}_{\text{McD}}$ , which corresponds to the assumptions of McDiarmid’s inequality [57], and is defined by

$$\mathcal{A}_{\text{McD}} := \left\{ (f, \mu) \left| \begin{array}{l} \mu = \mu_1 \otimes \mu_2 \otimes \mu_3, \\ m_1 = 5.5 \text{ mm}^2 \leq \mathbb{E}_\mu[f] \leq m_2 = 7.5 \text{ mm}^2, \\ \text{Osc}_i(f) \leq \text{Osc}_i(H) \text{ for } i = 1, 2, 3 \end{array} \right. \right\}. \tag{1.9}$$

**Definition 1.1.** Let  $\mathcal{X} := \mathcal{X}_1 \times \cdots \times \mathcal{X}_m$  and consider a function  $f: \mathcal{X} \rightarrow \mathbb{R}$ . For  $i = 1, \dots, m$ , we define the component-wise oscillations

$$\text{Osc}_i(f) := \sup_{(x_1, \dots, x_m) \in \mathcal{X}} \sup_{x'_i \in \mathcal{X}_i} |f(\dots, x_i, \dots) - f(\dots, x'_i, \dots)|. \tag{1.10}$$

Thus,  $\text{Osc}_i(f)$  measures the maximum oscillation of  $f$  in the  $i^{\text{th}}$  factor.

**Remark 1.2.** The explicit expression (1.5) of  $H$  and the ranges (1.7) allow us to compute the component-wise oscillations  $\text{Osc}_i(H)$ , which are, respectively,  $8.86 \text{ mm}^2$ ,  $4.17 \text{ mm}^2$  and  $7.20 \text{ mm}^2$  for thickness, obliquity, and velocity.



In general, for any admissible set  $\mathcal{A}$  of function/measure pairs for the perforation problem, we define

$$\mathcal{U}(\mathcal{A}) := \sup_{(f, \mu) \in \mathcal{A}} \mu[f(h, \theta, v) \leq 0]. \quad (1.11)$$

In this notation, the optimal upper bounds on the probability of non-perforation, given the information contained in  $\mathcal{A}_H$  and  $\mathcal{A}_{\text{McD}}$ , are  $\mathcal{U}(\mathcal{A}_H)$  and  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  respectively.

In  $\mathcal{A}_H$  the response function is exactly known whereas in  $\mathcal{A}_{\text{McD}}$  it is imperfectly known (the information on the response function is limited to its component-wise oscillations  $\text{Osc}_i(H)$ ). Both  $\mathcal{A}_H$  and  $\mathcal{A}_{\text{McD}}$  describe epistemic uncertainties (since in  $\mathcal{A}_H$  the probability distributions of thickness, obliquity, and velocity are imperfectly known).  $\mathcal{A}_{\text{McD}}$  is the set of response functions  $f$  and probability measures  $\mu$  that could be  $H$  and  $\mathbb{P}$  given the information contained in (i.e. the constraints imposed by)  $\text{Osc}_i(H)$ , the independence of the input variables and the bounds  $m_1$  and  $m_2$  on the mean perforation area.  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  quantifies the worst case scenario, i.e. the largest probability of non-perforation given what  $H$  and  $\mathbb{P}$  could be.

**Reduction theorems.** The optimization variables associated with  $\mathcal{U}(\mathcal{A}_H)$  are tensorizations of probability measures on thickness  $h$ , on obliquity  $\theta$  and velocity  $v$ . This problem is not directly computational tractable since finding the optimum appears to require a search over the spaces of probability measures on the intervals  $[1.524, 2.667]$  mm,  $[0, \frac{\pi}{6}]$  and  $[2.1, 2.8]$  km  $\cdot$  s $^{-1}$ . However, in Section 4 (Theorem 4.1 and Corollary 4.4) we show that, since the constraint  $m_1 \leq \mathbb{E}_\mu[H] \leq m_2$  is multi-linear in  $\mu_1, \mu_2$  and  $\mu_3$ , the optimum  $\mathcal{U}(\mathcal{A}_H)$  can be achieved by searching among those measures  $\mu$  whose marginal distributions on each of the three input parameter ranges have support on at most two points. That is,

$$\mathcal{U}(\mathcal{A}_H) = \mathcal{U}(\mathcal{A}_\Delta), \quad (1.12)$$

where the reduced feasible set  $\mathcal{A}_\Delta$  is given by

$$\mathcal{A}_\Delta := \left\{ (H, \mu) \left| \begin{array}{l} H \text{ given by (1.5),} \\ \mu = \mu_1 \otimes \mu_2 \otimes \mu_3, \\ \mu_i \in \Delta_1(\mathcal{X}_i) \text{ for } i = 1, 2, 3, \\ m_1 \leq \mathbb{E}_\mu[H] \leq m_2 \end{array} \right. \right\}, \quad (1.13)$$

where

$$\Delta_1(\mathcal{X}_i) := \{ \alpha \delta_{x^0} + (1 - \alpha) \delta_{x^1} \mid x^j \in \mathcal{X}_i, \text{ for } j = 0, 1 \text{ and } \alpha \in [0, 1] \}$$

denotes the set of binary convex combinations of Dirac masses on  $\mathcal{X}_i$ .

More generally, although the OUQ optimization problems (1.2) are extremely large, we show in Section 4 that an important subclass enjoys significant and practical finite-dimensional reduction properties. More precisely, although the optimization variables  $(f, \mu)$  live in a product space of functions and probability measures, for OUQ problems governed by linear inequality constraints on generalized moments, we demonstrate in Theorem 4.1 and Corollary 4.4 that the search can be reduced to one over probability measures that are products of finite convex combinations of Dirac masses with explicit

upper bounds on the number of Dirac masses. Moreover, all the results in this paper can be extended to sets of extreme points (extremal measures) more general than Dirac masses, such as those described by Dynkin [24]; we have phrased the results in terms of Dirac masses for simplicity.

Furthermore, when all constraints are generalized moments of functions of  $f$ , the search over admissible functions reduces to a search over functions on an  $m$ -fold product of finite discrete spaces, and the search over  $m$ -fold products of finite convex combinations of Dirac masses reduce to the products of probability measures on this  $m$ -fold product of finite discrete spaces. This latter reduction, presented in Theorem 4.7, completely eliminates dependency on the coordinate positions of the Dirac masses. With this result, the optimization variables of  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  can be reduced to functions and products of probability measures on  $\{0, 1\}^3$ .

### 1.2.2 Optimal concentration inequalities.

Concentration-of-measure inequalities can be used to obtain upper bounds on  $\mathcal{U}(\mathcal{A}_H)$  and  $\mathcal{U}(\mathcal{A}_{\text{McD}})$ ; in that sense, they lead to sub-optimal methods. Indeed, according to McDiarmid's inequality [57, 58], for all functions  $f$  of  $m$  independent variables, one must have

$$\mu[f(X_1, \dots, X_m) - \mathbb{E}_\mu[f] \geq a] \leq \exp\left(-2 \frac{a^2}{\sum_{i=1}^m (\text{Osc}_i(f))^2}\right). \quad (1.14)$$

Application of this inequality to (1.9) (using  $\mathbb{E}_\mu[f] \geq m_1 = 5.5 \text{ mm}^2$ ) yields the bound

$$\mathcal{U}(\mathcal{A}_{\text{McD}}) \leq \exp\left(-\frac{2m_1^2}{\sum_{i=1}^3 \text{Osc}_i(H)^2}\right) = 66.4\%. \quad (1.15)$$

Note that  $\mathcal{U}(\mathcal{A}_{\text{McD}}) := \sup_{(f, \mu) \in \mathcal{A}_{\text{McD}}} \mu[f \leq 0]$  is the least upper bound on the probability of non-perforation  $\mathbb{P}[H = 0]$  given the information contained in the admissible set (1.9).

In Section 5, the reduction techniques of Section 4 are applied to obtain optimal McDiarmid and Hoeffding inequalities, i.e. optimal concentration-of-measure inequalities with the assumptions of McDiarmid's inequality [57] or Hoeffding's inequality [35]. In particular, Theorems 5.1, 5.2 and 5.4 provide analytic solutions to the McDiarmid problem for dimension  $m = 1, 2, 3$ , and Proposition 5.7 provides a recursive formula for general  $m$ , thereby providing an optimal McDiarmid inequality in these cases. In Theorems 5.11 and 5.13, we give analytic solutions under Hoeffding's assumptions. A noteworthy result is that the optimal bounds associated with McDiarmid's and Hoeffding's assumptions are the same for  $m = 2$  but may be distinct for  $m = 3$ , and so, in some sense, information about linearity or non-linearity of the response function has a different effect depending upon the dimension  $m$  of the problem.

**Non-propagation of uncertainties.** For  $m = 2$ , define  $\mathcal{A}_2$  to be the space of all functions  $f$  and measure  $\mu$  such that  $\mu = \mu_1 \otimes \mu_2$  and  $\text{Osc}_i(f) \leq D_i$ . The optimal

concentration-of-measure inequality with the assumptions of McDiarmid’s inequality, Theorem 5.2, states that

$$\sup_{(f,\mu) \in \mathcal{A}_2} \mu[f(X_1, X_2) - \mathbb{E}_\mu[f] \geq a] = \begin{cases} 0, & \text{if } D_1 + D_2 \geq a, \\ \frac{(D_1 + D_2 - a)^2}{4D_1D_2}, & \text{if } |D_1 - D_2| \leq a \leq D_1 + D_2, \\ 1 - \frac{a}{\max(D_1, D_2)}, & \text{if } 0 \leq a \leq |D_1 - D_2|. \end{cases} \quad (1.16)$$

Observe that if  $D_2 + a \leq D_1$ , then the optimal bound does not depend on  $D_2$ , and therefore, any decrease in  $D_2$  does not improve the inequality. These explicit bounds show that, although uncertainties may propagate for the true values of  $G$  and  $\mathbb{P}$  (as expected from the sensitivity analysis paradigm), they may fail to do so when the information is incomplete on  $G$  and  $\mathbb{P}$  and the objective is the maximum of  $\mu[f \geq a]$  compatible with the given information. The non-propagation of input uncertainties is a non-trivial observation related to the fact that some of the constraints defining the range of the input variables may not be realized by the worst-case scenario (extremum of the OUQ problem). We have further illustrated this point in Section 8 and shown that for systems characterized by multiple scales or hierarchical structures, information or uncertainties may not propagate across scales. Note that the  $m = 2$  case does not apply to the SPHIR example (since (1.9) involves three variables, i.e. it requires  $m = 3$ ).

**Application to the SPHIR facility admissible set (1.9).** For  $m = 3$ , the “optimal McDiarmid inequality” of Theorem 5.4 and Remark 4.2 provides the upper bound

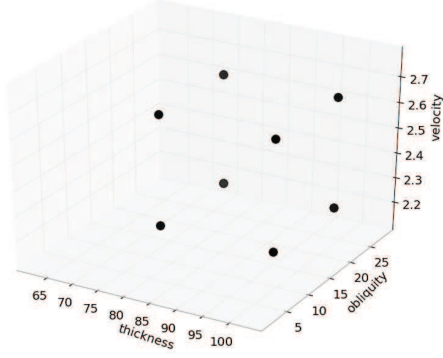
$$\mathcal{U}(\mathcal{A}_{\text{McD}}) = 43.7\%. \quad (1.17)$$

Remark 5.6 also shows that reducing the uncertainty in obliquity (described by the constraint  $\text{Osc}_i(f) \leq \text{Osc}_i(H)$  in (1.9) for  $i = \text{obliquity}$ ) does not affect the *least* upper bound (1.17). Recall that  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  is the least upper bound on the probability that the perforation is zero given that the mean perforation area is in between  $5.5 \text{ mm}^2$  and  $7.5 \text{ mm}^2$  and the constraints imposed by the independence, ranges and component-wise oscillations associated with the input random variables.

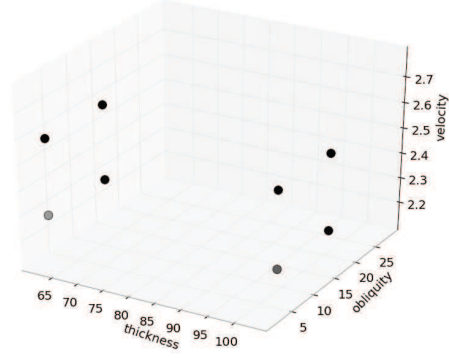
The difference between (1.15) and (1.17) lies in the fact that 66.4% is non-optimal whereas 43.7% is the *least* upper bound on the probability of non perforation given the information contained in  $\mathcal{A}_{\text{McD}}$ . 43.7% is a direct function of the information contained in  $\mathcal{A}_{\text{McD}}$  and Section 2 shows how admissible sets with higher information content lead to smaller least upper bounds on the probability of non perforation. Section 2 also shows how such admissible sets can be constructed, in the OUQ framework, via the optimal selection of experiments.

### 1.2.3 Computational framework.

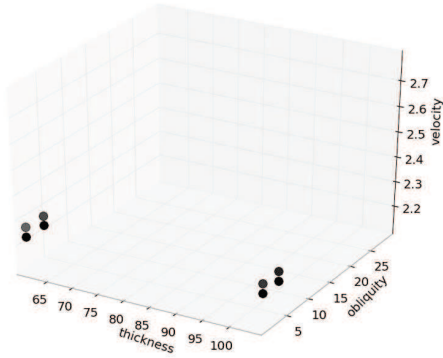
With access to  $H$ , not just its componentwise oscillations, even sharper bounds on the probability of non-perforation can be calculated. Although we do not have an analytical



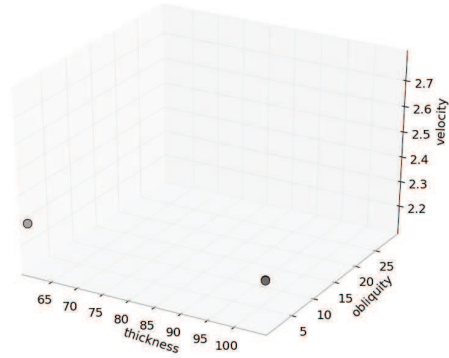
(a) support points at iteration 0



(b) support points at iteration 150



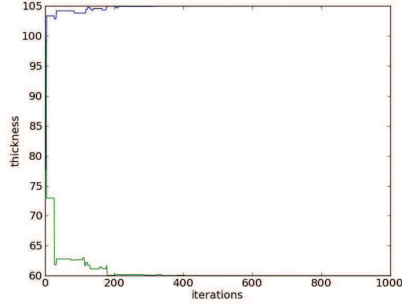
(c) support points at iteration 200



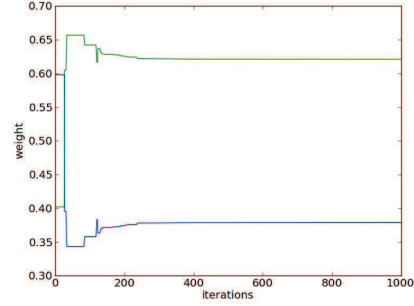
(d) support points at iteration 1000

Figure 1.3: For  $\#\text{supp}(\mu_i) \leq 2$ ,  $i = 1, 2, 3$ , the maximizers of the OUQ problem (1.12) associated with the information set (1.8) collapse to two-point (as opposed to eight-point) support. Velocity and obliquity marginals each collapse to a single Dirac mass, while the plate thickness marginal collapses to have support on the extremes of its range. Note the perhaps surprising result that the probability of non-perforation is maximized by a distribution supported on the minimal, not maximal, impact obliquity.

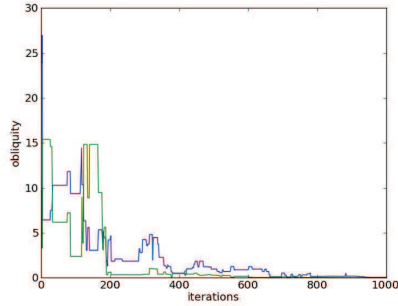
Figure 1.4: Time evolution of the genetic algorithm search for the OUQ problem (1.12) associated with the information set (1.8) ((1.13) after reduction) for  $\#\text{supp}(\mu_i) \leq 2$  for  $i = 1, 2, 3$ , as optimized by *mystic*. Thickness quickly converges to the extremes of its range, with a weight of 0.621 at 60 mils and a weight of 0.379 at 105 mils. The degeneracy in obliquity at 0 causes the fluctuations seen in the convergence of obliquity weight. Similarly, velocity converges to a single support point at  $2.289 \text{ km} \cdot \text{s}^{-1}$ , the ballistic limit velocity for thickness 105 mils and obliquity 0 (see (1.6)).



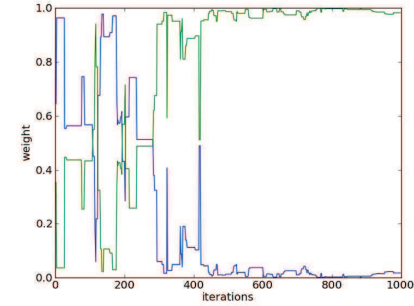
(a) convergence for thickness



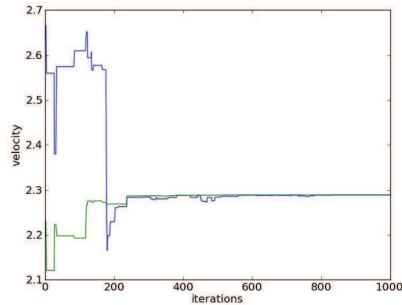
(b) convergence for thickness weight



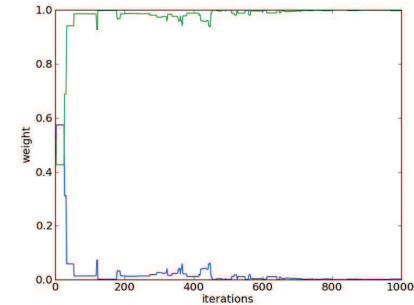
(c) convergence for obliquity



(d) convergence for obliquity weight



(e) convergence for velocity



(f) convergence for velocity weight

formula for  $\mathcal{U}(\mathcal{A}_H)$ , its calculation is made possible by the identity (1.12) derived from the reduction results of Section 4. A numerical optimization over the finite-dimensional reduced feasible set  $\mathcal{A}_\Delta$  using a Differential Evolution [72] optimization algorithm implemented in the *mystic framework* [60] (see Subsection 6.3) yields the following optimal upper bound on the probability of non-perforation:

$$\mathbb{P}[H = 0] \leq \mathcal{U}(\mathcal{A}_H) = \mathcal{U}(\mathcal{A}_\Delta) \stackrel{\text{num}}{=} 37.9\%.$$

Observe that “ $\mathbb{P}[H = 0] \leq \mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_\Delta)$ ” is a theorem, whereas “ $\mathcal{U}(\mathcal{A}_\Delta) \stackrel{\text{num}}{=} 37.9\%$ ” is the output of an algorithm (in this case, a genetic algorithm implemented in the *mystic framework* described in Subsection 6.3). In particular, its validity is correlated with the efficiency of the specific algorithm. We have introduced the symbol  $\stackrel{\text{num}}{=}$  to emphasize the distinction between mathematical (in)equalities and numerical outputs.

Although we do not have a theorem associated with the convergence of the numerical optimization algorithm, we have a robust control over its efficiency because it is applied to the finite dimensional problem  $\mathcal{U}(\mathcal{A}_\Delta)$  instead of the infinite optimization problem associated with  $\mathcal{U}(\mathcal{A}_H)$  (thanks to the reduction theorems obtained in Section 4).

We also observe that the maximizer  $\mathcal{U}(\mathcal{A}_H)$  can be of significantly smaller dimension than that of the elements of  $\mathcal{U}(\mathcal{A}_\Delta)$ . Indeed, for  $\#\text{supp}(\mu_i) \leq 2$ ,  $i = 1, 2, 3$  (where  $\#\text{supp}(\mu_i)$  is the number of points forming the support of  $\mu_i$ ), Figure 1.3 shows that numerical simulations collapse to two-point support: the velocity and obliquity marginals each collapse to a single Dirac mass, and the plate thickness marginal collapses to have support on the two extremes of its range. See Figure 1.4 for plots of the locations and weights of the Dirac masses forming each marginal  $\mu_i$  as functions of the number of iterations. Note that the lines for *thickness* and *thickness weight* are of the same color if they correspond to the same support point for the measure.

In Section 6 we observe that, even when a wider search is performed (i.e. over measures with more than two-point support per marginal), the calculated maximizers for these problems maintain two-point support. This observation suggests that the extreme points of the reduced OUQ problems are, in some sense, attractors and that adequate numerical implementation of OUQ problems can detect and use “hidden” reduction properties even in the absence of theorems proving them to be true. Based on these observations, we propose, in Section 6, an OUQ optimization algorithm for arbitrary constraints based on a coagulation/fragmentation of probability distributions.

The simulations of Figures 1.3 and 1.4 show that extremizers are singular and that their support points identify key players, i.e. weak points of the system. In particular, for  $\mathcal{U}(\mathcal{A}_H)$ , the location of the two-point support extremizer shows that reducing maximum obliquity and the range of velocity will not decrease the optimal bound on the probability of non perforation, and suggests that reducing the uncertainty in thickness will decrease this bound. In Section 2, we will show that the OUQ framework allows the development of an OUQ loop that can be used for experimental design. In particular, we will show that the problem of predicting optimal bounds on the results of experiments under the assumption that the system is safe (or unsafe) is well-posed and benefits from similar reduction properties as the certification problem. Best experiments are then naturally

identified as those whose predicted ranges have minimal overlap between safe and unsafe systems.

#### 1.2.4 Outline of the paper.

Section 2 formally describes the *Optimal Uncertainty Quantification* framework and what it means to give optimal bounds on the probability of failure in (1.1) given (limited) information/assumptions about the system of interest, and hence how to rigorously certify or de-certify that system. Section 3 shows that many important UQ problems, such as prediction, verification and validation, can be formulated as certification problems. Section 3 also gives a comparison of OUQ with other widely used UQ methods. Section 4 shows that although OUQ optimization problems (1.2) are (a priori) infinite-dimensional, they can (in general) be reduced to equivalent finite-dimensional problems in which the optimization is over the extremal scenarios of  $\mathcal{A}$  and that the dimension of the reduced problem is a function of the number of probabilistic inequalities that describe  $\mathcal{A}$ . Just as a linear program finds its extreme value at the extremal points of a convex domain in  $\mathbb{R}^n$ , OUQ problems reduce to searches over finite-dimensional families of extremal scenarios. Section 5 applies the results of Section 4 to obtain optimal concentration inequalities under the assumptions of McDiarmid’s inequality and Hoeffding’s inequality. Those inequalities show that, although uncertainties may propagate for the true values of  $G$  and  $\mathbb{P}$ , they might not when the information is incomplete on  $G$  and  $\mathbb{P}$ . Section 6 discusses the numerical implementation of OUQ algorithms for the analytical surrogate model (1.5) for hypervelocity impact. Section 7 assesses the feasibility of the OUQ formalism by means of an application to the safety assessment of truss structures subjected to ground motion excitation. This application contains many of the features that both motivate and challenge UQ, including imperfect knowledge of random inputs of high dimensionality, a time-dependent and complex response of the system, and the need to make high-consequence decisions pertaining to the safety of the system. Section 8 applies the OUQ framework and reduction theorems of sections 4 and 5 to divergence form elliptic PDEs. A striking observation of Section 8 is that with incomplete information on the probability distribution of the microstructure, uncertainties or information do not necessarily propagate across scales. Section 9 emphasizes that the “UQ problem” (as it is treated in common practice today) appears to have all the symptoms of an ill-posed problem; and that, at the very least, it lacks a coherent general presentation, much like the state of probability theory before its rigorous formulation by Kolmogorov in the 1930s. It also stresses that OUQ is not the definitive answer to the UQ problem, but an opening gambit. Section 10 gives the proofs of our main results.

## 2 Optimal Uncertainty Quantification

In this section, we describe more formally the *Optimal Uncertainty Quantification* framework. In particular, we describe what it means to give optimal bounds on the probability of failure in (1.1) given information/assumptions about the system of interest, and hence

how to rigorously certify or de-certify that system.

For the sake of clarity, we will start the description of OUQ with deterministic information and assumptions (when  $\mathcal{A}$  is a deterministic set of functions and probability measures).

## 2.1 First description

In the OUQ paradigm, information and assumptions lie at the core of UQ: the available information and assumptions describe sets of admissible scenarios over which optimizations will be performed. As noted by Hoeffding [34], assumptions about the system of interest play a central and sensitive role in any statistical decision problem, even though the assumptions are often only approximations of reality.

A simple example of an information/assumptions set is given by constraining the mean and range of the response function. For example, let  $\mathcal{M}(\mathcal{X})$  be the set of probability measures on the set  $\mathcal{X}$ , and let  $\mathcal{A}_1$  be the set of pairs of probability measures  $\mu \in \mathcal{M}(\mathcal{X})$  and real-valued measurable functions  $f$  on  $\mathcal{X}$  such that the mean value of  $f$  with respect to  $\mu$  is  $b$  and the diameter of the range of  $f$  is at most  $D$ ;

$$\mathcal{A}_1 := \left\{ (f, \mu) \left| \begin{array}{l} f: \mathcal{X} \rightarrow \mathbb{R}, \\ \mu \in \mathcal{M}(\mathcal{X}), \\ \mathbb{E}_\mu[f] = b, \\ (\sup f - \inf f) \leq D \end{array} \right. \right\}. \quad (2.1)$$

Let us assume that *all* that we know about the “reality”  $(G, \mathbb{P})$  is that  $(G, \mathbb{P}) \in \mathcal{A}_1$ . Then any other pair  $(f, \mu) \in \mathcal{A}_1$  constitutes an admissible scenario representing a valid possibility for the “reality”  $(G, \mathbb{P})$ . If asked to bound  $\mathbb{P}[G(X) \geq a]$ , should we apply different methods and obtain different bounds on  $\mathbb{P}[G(X) \geq a]$ ? Since some methods will distort this information set and others are only using part of it, we instead view set  $\mathcal{A}_1$  as a feasible set for an optimization problem.

**The general OUQ framework.** In the general case, we regard the response function  $G$  as an unknown measurable function, with some possibly known characteristics, from one measurable space  $\mathcal{X}$  of inputs to a second measurable space  $\mathcal{Y}$  of values. The input variables are generated randomly according to an unknown random variable  $X$  with values in  $\mathcal{X}$  according to a law  $\mathbb{P} \in \mathcal{M}(\mathcal{X})$ , also with some possibly known characteristics. We let a measurable subset  $\mathcal{Y}_0 \subseteq \mathcal{Y}$  define the *failure region*; in the example given above,  $\mathcal{Y} = \mathbb{R}$  and  $\mathcal{Y}_0 = [a, +\infty)$ . When there is no danger of confusion, we shall simply write  $[G \text{ fails}]$  for the event  $[G(X) \in \mathcal{Y}_0]$ .

Let  $\epsilon \in [0, 1]$  denote the *greatest acceptable probability of failure*. We say that the system is *safe* if  $\mathbb{P}[G \text{ fails}] \leq \epsilon$  and the system is *unsafe* if  $\mathbb{P}[G \text{ fails}] > \epsilon$ . By *information*, or a *set of assumptions*, we mean a subset

$$\mathcal{A} \subseteq \left\{ (f, \mu) \left| \begin{array}{l} f: \mathcal{X} \rightarrow \mathcal{Y} \text{ is measurable,} \\ \mu \in \mathcal{M}(\mathcal{X}) \end{array} \right. \right\} \quad (2.2)$$



that contains, at the least,  $(G, \mathbb{P})$ . The set  $\mathcal{A}$  encodes all the information that we have about the real system  $(G, \mathbb{P})$ , information that may come from known physical laws, past experimental data, and expert opinion. In the example  $\mathcal{A}_1$  above, the only information that we have is that the mean response of the system is  $b$  and that the diameter of its range is at most  $D$ ; any pair  $(f, \mu)$  that satisfies these two criteria is an *admissible scenario* for the unknown reality  $(G, \mathbb{P})$ . Since some admissible scenarios may be safe (i.e. have  $\mu[f \text{ fails}] \leq \epsilon$ ) whereas other admissible scenarios may be unsafe (i.e. have  $\mu[f \text{ fails}] > \epsilon$ ), we decompose  $\mathcal{A}$  into the disjoint union  $\mathcal{A} = \mathcal{A}_{\text{safe}, \epsilon} \uplus \mathcal{A}_{\text{unsafe}, \epsilon}$ , where

$$\mathcal{A}_{\text{safe}, \epsilon} := \{(f, \mu) \in \mathcal{A} \mid \mu[f \text{ fails}] \leq \epsilon\}, \quad (2.3a)$$

$$\mathcal{A}_{\text{unsafe}, \epsilon} := \{(f, \mu) \in \mathcal{A} \mid \mu[f \text{ fails}] > \epsilon\}. \quad (2.3b)$$

Now observe that, given such an information/assumptions set  $\mathcal{A}$ , there exist upper and lower bounds on  $\mathbb{P}[G(X) \geq a]$  corresponding to the scenarios compatible with assumptions, i.e. the values  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$  of the optimization problems:

$$\mathcal{L}(\mathcal{A}) := \inf_{(f, \mu) \in \mathcal{A}} \mu[f \text{ fails}] \quad (2.4a)$$

$$\mathcal{U}(\mathcal{A}) := \sup_{(f, \mu) \in \mathcal{A}} \mu[f \text{ fails}]. \quad (2.4b)$$

Since  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$  are well-defined in  $[0, 1]$ , and approximations are sufficient for most purposes and are necessary in general, the difference between sup and max should not be much of an issue. Of course, some of the work that follows is concerned with the attainment of maximizers, and whether those maximizers have any simple structure that can be exploited for the sake of computational efficiency, and this is the topic of Section 4. For the moment, however, simply assume that  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$  can indeed be computed on demand. Now, since  $(G, \mathbb{P}) \in \mathcal{A}$ , it follows that

$$\mathcal{L}(\mathcal{A}) \leq \mathbb{P}[G \text{ fails}] \leq \mathcal{U}(\mathcal{A}).$$

Moreover, the upper bound  $\mathcal{U}(\mathcal{A})$  is *optimal* in the sense that

$$\mu[f \text{ fails}] \leq \mathcal{U}(\mathcal{A}) \text{ for all } (f, \mu) \in \mathcal{A}$$

and, if  $U' < \mathcal{U}(\mathcal{A})$ , then there is an admissible scenario  $(f, \mu) \in \mathcal{A}$  such that

$$U' < \mu[f \text{ fails}] \leq \mathcal{U}(\mathcal{A}).$$

That is, although  $\mathbb{P}[G \text{ fails}]$  may be much smaller than  $\mathcal{U}(\mathcal{A})$ , there is a pair  $(f, \mu)$  which satisfies the same assumptions as  $(G, \mathbb{P})$  such that  $\mu[f \text{ fails}]$  is approximately equal to  $\mathcal{U}(\mathcal{A})$ . Similar remarks apply for the lower bound  $\mathcal{L}(\mathcal{A})$ .

Moreover, the values  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$ , defined in (2.4) can be used to construct a solution to the certification problem. Let the certification problem be defined by an error function that gives an error whenever 1) the certification process produces “safe” and there exists an admissible scenario that is unsafe, 2) the certification process produces

	$\mathcal{L}(\mathcal{A}) := \inf_{(f,\mu) \in \mathcal{A}} \mu[f(X) \geq a]$	$\mathcal{U}(\mathcal{A}) := \sup_{(f,\mu) \in \mathcal{A}} \mu[f(X) \geq a]$
$\leq \epsilon$	<b>Cannot decide</b> Insufficient Information	<b>Certify</b> Safe even in the Worst Case
$> \epsilon$	<b>De-certify</b> <b>Unsafe</b> even in the Best Case	<b>Cannot decide</b> Insufficient Information

Table 2.1: The OUQ certification process provides a rigorous certification criterion whose outcomes are of three types: “Certify”, “De-certify” and “Cannot decide”.

“unsafe” and there exists an admissible scenario that is safe, or 3) the certification process produces “cannot decide” and all admissible scenarios are safe or all admissible points are unsafe; otherwise, the certification process produces no error. The following proposition demonstrates that, except in the special case  $\mathcal{L}(\mathcal{A}) = \epsilon$ , that these values determine an optimal solution to this certification problem.

**Proposition 2.1.** *If  $(G, \mathbb{P}) \in \mathcal{A}$  and*

- $\mathcal{U}(\mathcal{A}) \leq \epsilon$  then  $\mathbb{P}[G \text{ fails}] \leq \epsilon$ .
- $\epsilon < \mathcal{L}(\mathcal{A})$  then  $\mathbb{P}[G \text{ fails}] > \epsilon$ .
- $\mathcal{L}(\mathcal{A}) < \epsilon < \mathcal{U}(\mathcal{A})$  then there exists  $(f_1, \mu_1) \in \mathcal{A}$  and  $(f_2, \mu_2) \in \mathcal{A}$  such that  $\mu_1[f_1 \text{ fails}] < \epsilon < \mu_2[f_2 \text{ fails}]$ .

In other words, provided that the information set  $\mathcal{A}$  is valid (in the sense that  $(G, \mathbb{P}) \in \mathcal{A}$ ) then if  $\mathcal{U}(\mathcal{A}) \leq \epsilon$ , then, the system is provably safe; if  $\epsilon < \mathcal{L}(\mathcal{A})$ , then the system is provably unsafe; and if  $\mathcal{L}(\mathcal{A}) < \epsilon < \mathcal{U}(\mathcal{A})$ , then the safety of the system cannot be decided due to lack of information. The corresponding certification process and its optimality are represented in Table 2.1. Hence, solving the optimization problems (2.4) determines an optimal solution to the certification problem, under the condition that  $\mathcal{L}(\mathcal{A}) \neq \epsilon$ . When  $\mathcal{L}(\mathcal{A}) = \epsilon$  we can still produce an optimal solution if we obtain further information. That is, when  $\mathcal{L}(\mathcal{A}) = \epsilon = \mathcal{U}(\mathcal{A})$ , then the optimal process produces “safe”. On the other hand, when  $\mathcal{L}(\mathcal{A}) = \epsilon < \mathcal{U}(\mathcal{A})$ , the optimal solution depends on whether or not there exists a minimizer  $(f, \mu) \in \mathcal{A}$  such that  $\mu[f \text{ fails}] = \mathcal{L}(\mathcal{A})$ ; if so, the optimal process should declare “cannot decide”, otherwise, the optimal process should declare “unsafe”. Observe that, in Table 2.1, we have classified  $\mathcal{L}(\mathcal{A}) = \epsilon < \mathcal{U}(\mathcal{A})$  as “cannot decide”. This “nearly optimal” solution appears natural and conservative without the knowledge of the existence or non-existence of optimizers.

**Example 2.1.** The bounds  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$  can be computed exactly — and are non-trivial — in the case of the simple example  $\mathcal{A}_1$  given in (2.1). Indeed, writing  $x_+ := \max(x, 0)$ , the optimal upper bound is given by

$$\mathcal{U}(\mathcal{A}_1) = p_{\max} := \left(1 - \frac{(a-b)_+}{D}\right)_+, \quad (2.5)$$

where the maximum is achieved by taking the measure of probability of the random variable  $f(X)$  to be the weighted sum of two weighted Dirac delta masses<sup>1</sup>

$$p_{\max}\delta_a + (1 - p_{\max})\delta_{a-D}.$$

This simple example demonstrates an extremely important point: even if the function  $G$  is extremely expensive to evaluate, certification can be accomplished without recourse to the expensive evaluations of  $G$ . Furthermore, the simple structure of the maximizers motivates the reduction theorems later in Section 4.

**Example 2.2.** As shown in Equation (1.14), concentration-of-measure inequalities lead to sub-optimal methods in the sense that they can be used to obtain upper bounds on  $\mathcal{U}(\mathcal{A})$  and lower bounds on  $\mathcal{L}(\mathcal{A})$ . Observe that McDiarmid's inequality (1.14) required an information/assumptions set  $\mathcal{A}_{\text{McD}}$  where the space  $\mathcal{X}$  is a product space with  $X = (X_1, X_2, \dots, X_m)$ , the mean performance satisfies  $\mathbb{E}[G(X)] \leq b$ , the  $m$  inputs  $X_1, \dots, X_m$  are independent, and the component-wise oscillations of  $G$ , (see (1.10)) are bounded  $\text{Osc}_i(G) \leq D_i$ . It follows from McDiarmid's inequality (1.14) that, under the assumptions  $\mathcal{A}_{\text{McD}}$ ,

$$\mathcal{U}(\mathcal{A}_{\text{McD}}) \leq \exp\left(-2\frac{(a-b)_+^2}{\sum_{i=1}^m D_i^2}\right).$$

This example shows how existing techniques such as concentration-of-measure inequalities can be incorporated into OUQ. In Section 4, we will show how to reduce  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  to a finite dimensional optimization problem. Based on this reduction, in Section 5, we provide analytic solutions to the optimization problem  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  for  $m = 1, 2, 3$ . In practice, the computation of the bounds  $D_i$  require the resolution of an optimization problem, we refer to [52, 46, 1] for practical methods. We refer to [52, 46, 1, 78] for illustrations of UQ through concentration of measure inequalities. In particular, since  $\text{Osc}_i(G)$  is a semi-norm, a (possibly numerical) model can be used to compute bounds on component-wise oscillations of  $G$  via the triangular inequality  $\text{Osc}_i(G) \leq \text{Osc}_i(F) + \text{Osc}_i(G - F)$  (we refer to [52, 46, 1] for details, the idea here is that an accurate model leads to a reduced number of experiments for the computation of  $\text{Osc}_i(G - F)$ , while the computation of  $\text{Osc}_i(F)$  does not involve experiments). In the sequel we will refer to  $D_{i,G} := \text{Osc}_i(G)$  (for  $i = 1, \dots, m$ ) as the *sub-diameters* of  $G$  and to

$$D_G := \sqrt{\sum_{i=1}^m D_{i,G}^2} \tag{2.6}$$

as the *diameter* of  $G$ . Bounds on  $\text{Osc}_i(G)$  are useful because they constitute a form of non-linear sensitivity analysis and, combined with independence constraints, they lead

---

<sup>1</sup> $\delta_z$  is the Dirac delta mass on  $z$ , i.e. the measure of probability on Borel subsets  $A \subset \mathbb{R}$  such that  $\delta_z(A) = 1$  if  $z \in A$  and  $\delta_z(A) = 0$  otherwise. The first Dirac delta mass is located at the minimum of the interval  $[a, \infty]$  (since we are interested in maximizing the probability of the event  $\mu[f(X) \geq a]$ ). The second Dirac delta mass is located at  $z = a - D$  because we seek to maximize  $p_{\max}$  under the constraints  $p_{\max}a + (1 - p_{\max})z \leq b$  and  $a - z \leq D$ .

Admissible scenarios, $\mathcal{A}$	$\mathcal{U}(\mathcal{A})$	Method
$\mathcal{A}_{\text{McD}}$ : independence, oscillation and mean constraints as given by (1.9)	$\leq 66.4\%$ $= 43.7\%$	McDiarmid's inequality Theorem 5.4
$\mathcal{A}_H$ as given by (1.8)	$\stackrel{\text{num}}{=} 37.9\%$	<i>mystic</i> , $H$ known
$\mathcal{A}_H \cap \left\{ (H, \mu) \mid \begin{array}{l} \mu\text{-median velocity} \\ = 2.45 \text{ km} \cdot \text{s}^{-1} \end{array} \right\}$	$\stackrel{\text{num}}{=} 30.0\%$	<i>mystic</i> , $H$ known
$\mathcal{A}_H \cap \left\{ (H, \mu) \mid \mu\text{-median obliquity} = \frac{\pi}{12} \right\}$	$\stackrel{\text{num}}{=} 36.5\%$	<i>mystic</i> , $H$ known
$\mathcal{A}_H \cap \left\{ (H, \mu) \mid \text{obliquity} = \frac{\pi}{6} \mu\text{-a.s.} \right\}$	$\stackrel{\text{num}}{=} 28.0\%$	<i>mystic</i> , $H$ known

Table 2.2: Summary of the upper bounds on the probability of non-perforation for Example (1.8) obtained by various methods and assumptions. Note that OUQ calculations using *mystic* (described in Section 6) involve evaluations of the function  $H$ , whereas McDiarmid's inequality and the optimal bound given the assumptions of McDiarmid's inequality use only the mean of  $H$  and its McDiarmid subdiameters, not  $H$  itself. Note also that the incorporation of additional information/assumptions, e.g. on impact obliquity, always reduces the OUQ upper bound on the probability of failure, since this corresponds to a restriction to a subset of the original feasible set  $\mathcal{A}_H$  for the optimization problem.

to the concentration of measure phenomenon. The OUQ methodology can also handle constraints of the type  $\|G - F\| < C$  (which are not sufficient to observe take advantage of the concentration of measure effect) and  $G(x_i) = z_i$  [90].

**Example 2.3.** For the set  $\mathcal{A}_H$  given in Equation (1.8), the inclusion of additional information further reduces the upper bound  $\mathcal{U}(\mathcal{A}_H)$ . Indeed, the addition of assumptions lead to a smaller admissible set  $\mathcal{A}_H \mapsto \mathcal{A}' \subset \mathcal{A}_H$ , therefore  $\mathcal{U}$  decreases and  $\mathcal{L}$  increases. For example, if the median of the third input parameter (velocity) is known to lie at the midpoint of its range, and this information is used to provide an additional constraint, then the least upper bound on the probability of non-perforation drops to 30.0%. See Table 2.2 for a summary of the bounds presented in the hypervelocity impact example introduced in Subsection 1.2, and for further examples of the effect of additional information/constraints. The bounds given in Table 2.2 have been obtained by using the reduction theorems of Section 4 and the computational framework described in Section 6.

**Remark 2.2.** The number of iterations and evaluations of  $H$  associated with Table 2.2 are: 600 iterations and 15300  $H$ -evaluations (second row), 822 iterations and 22700  $H$ -evaluations (third row), 515 iterations and 14550  $H$ -evaluations (fourth row), 760 iterations and 18000  $H$ -evaluations (fifth row). Half of these numbers of iterations are

usually sufficient to obtain the extrema with 4 digits of accuracy (for the third row, for instance, 365 iterations and 9350  $H$ -evaluations are sufficient to obtain the first 4 decimal points of the optimum).

**On the selectiveness of the information set  $\mathcal{A}$ .** Observe that, except for the bound obtained from McDiarmid’s inequality, the bounds obtained in Table 2.2 are the best possible given the information contained in  $\mathcal{A}$ . If the unknown distribution  $\mathbb{P}$  is completely specified, say by restricting to the feasible set  $\mathcal{A}_{\text{unif}}$  for which the only admissible measure is the uniform probability measure on the cube  $\mathcal{X}$  (in which case the mean perforation area is  $\mathbb{E}[H] = 6.58 \text{ mm}^2$ ), then the probability of non-perforation is  $\mathcal{U}(\mathcal{A}_{\text{unif}}) = \mathcal{L}(\mathcal{A}_{\text{unif}}) \stackrel{\text{num}}{=} 3.8\%$ . One may argue that there is a large gap between the fifth (28%) row of Table 2.2 and 3.8% but observe that 3.8% relies on the exact knowledge of  $G$  (called  $H$  here) and  $\mathbb{P}$  whereas 28% relies on the limited knowledge contained in  $\mathcal{A}_H \cap \{(H, \mu) \mid \text{obliquity} = \frac{\pi}{6} \mu\text{-a.s.}\}$  with respect to which 28% is optimal. In particular, the gap between those two values is not caused by a lack of tightness of the method, but by a lack of selectiveness of the information contained in  $\mathcal{A}_H \cap \{(H, \mu) \mid \text{obliquity} = \frac{\pi}{6} \mu\text{-a.s.}\}$ . The (mis)use of the terms “tight” and “sharp” without reference to available information (and in presence of asymmetric information) can be the source of much confusion, something that we hope is cleared up by the present work. Given prior knowledge of  $G$  and  $\mathbb{P}$ , it would be an easy task to construct a set  $\mathcal{A}_{\mathbb{P}, G}$  containing  $(G, \mathbb{P})$  such that  $\mathcal{U}(\mathcal{A}_{\mathbb{P}, G}) \approx 4\%$  (if the probability of failure under  $(G, \mathbb{P})$  is 3.8%), but doing so would be delaying a honest discussion on one of the issues at the core of UQ: *How to construct  $\mathcal{A}$  without prior knowledge of  $G$  and  $\mathbb{P}$ ?* In other words, how to improve the “selectiveness” of  $\mathcal{A}$  or how to design experiments leading to “narrow”  $\mathcal{A}$ s? We will now show how this question can be addressed within the OUQ framework.

## 2.2 The Optimal UQ loop

In the previous subsection we discussed how the basic inequality

$$\mathcal{L}(\mathcal{A}) \leq \mathbb{P}[G \geq a] \leq \mathcal{U}(\mathcal{A})$$

provides rigorous optimal certification criteria. The certification process should not be confused with its three possible outcomes (see Table 2.1) which we call “certify” (we assert that the system is safe), “de-certify” (we assert that the system is unsafe) and “cannot decide” (the safety or un-safety of the system is undecidable given the information/assumption set  $\mathcal{A}$ ). Indeed, in the case

$$\mathcal{L}(\mathcal{A}) \leq \epsilon < \mathcal{U}(\mathcal{A})$$

there exist admissible scenarios under which the system is safe, and other admissible scenarios under which it is unsafe. Consequently, it follows that we can make no definite certification statement for  $(G, \mathbb{P})$  without introducing further information/assumptions. If no further information can be obtained, we conclude that we “cannot decide” (this

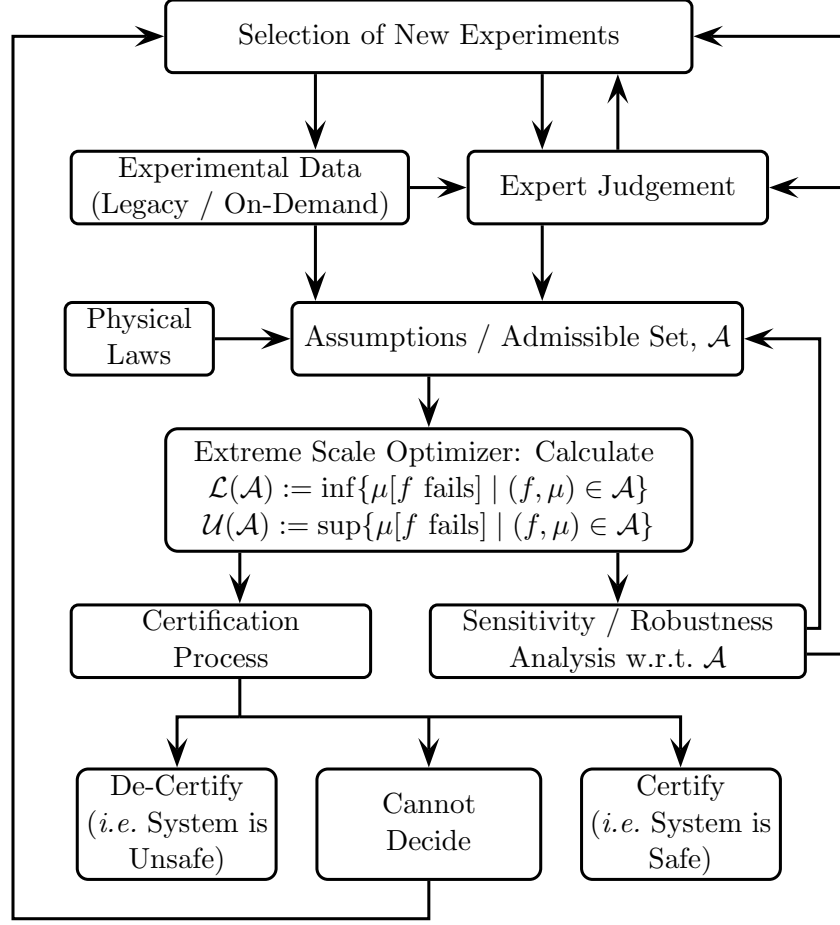


Figure 2.1: Optimal Uncertainty Quantification Loop.

state could also be called “do not decide”, because we could (arbitrarily) decide that the system is unsafe due to lack of information, for instance, but do not).

However, if sufficient resources exist to gather additional information, then we enter what may be called the *optimal uncertainty quantification loop*, illustrated in Figure 2.1. The admissible set  $\mathcal{A}$  draws on three principal sources of information: known physical laws, expert opinion, and experimental data. Once the set  $\mathcal{A}$  has been constructed, the calculation of the lower and upper bounds  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$  are well-posed optimization problems. If the results of these optimization problems lead to certification or de-certification, then we are done; if not, then new experiments should be designed and expert opinion sought in order to validate or invalidate the extremal scenarios that cause the inequality

$$\mathcal{L}(\mathcal{A}) \leq \epsilon < \mathcal{U}(\mathcal{A})$$

to hold. The addition of information means further constraints on the collection of

admissible scenarios; that is, the original admissible set  $\mathcal{A}$  is reduced to a smaller one  $\mathcal{A}' \subset \mathcal{A}$ , thereby providing sharper bounds on the probability of failure:

$$\mathcal{L}(\mathcal{A}) \leq \mathcal{L}(\mathcal{A}') \leq \mathbb{P}[G(X) \geq a] \leq \mathcal{U}(\mathcal{A}') \leq \mathcal{U}(\mathcal{A}).$$

The sharper bounds may meet the “certify” or “decertify” criteria of Table 2.1. If not, and there are still resources for gathering additional information, then the loop should be repeated. This process is the feedback arrow on the left-hand side of Figure 2.1.

The right-hand side of Figure 2.1 constitutes another aspect of the OUQ loop. The bounds  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$  are only useful insofar as the assumptions  $\mathcal{A}$  are accurate. It is possible that the sources of information that informed  $\mathcal{A}$  may have been in error: physical laws may have been extended outside their range of validity (e.g. Newtonian physics may have been applied in the relativistic regime), there may have been difficulties with the experiments or the results misinterpreted, or expert opinion may have been erroneous. Therefore, a vital part of OUQ is to examine the sensitivity and robustness of the bounds  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$  with respect to the assumption set  $\mathcal{A}$ . If the bounds  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$  are found to depend sensitively on one particular assumption (say, the mean performance of one component of the system), then it would be advisable to expend resources investigating this assumption.

The loop illustrated in Figure 2.1 differs from the loop used to solve the numerical optimization problem as described in Sub-section 6.3 and Remark 6.3.

**Experimental design and selection of the most decisive experiment.** An important aspect of the OUQ loop is the selection of new experiments. Suppose that a number of possible experiments  $E_i$  are proposed, each of which will determine some functional  $\Phi_i(G, \mathbb{P})$  of  $G$  and  $\mathbb{P}$ . For example,  $\Phi_1(G, \mathbb{P})$  could be  $\mathbb{E}_{\mathbb{P}}[G]$ ,  $\Phi_2(G, \mathbb{P})$  could be  $\mathbb{P}[X \in A]$  for some subset  $A \subseteq \mathcal{X}$  of the input parameter space, and so on. Suppose that there are insufficient experimental resources to run all of these proposed experiments. Let us now consider which experiment should be run for the certification problem. Recall that the admissible set  $\mathcal{A}$  is partitioned into safe and unsafe subsets as in (2.3). Define  $J_{\text{safe}, \epsilon}(\Phi_i)$  to be the closed interval spanned by the possible values for the functional  $\Phi_i$  over the safe admissible scenarios (i.e. the closed convex hull of the range of  $\Phi_i$  on  $\mathcal{A}_{\text{safe}, \epsilon}$ ): that is, let

$$J_{\text{safe}, \epsilon}(\Phi_i) := \left[ \inf_{(f, \mu) \in \mathcal{A}_{\text{safe}, \epsilon}} \Phi_i(f, \mu), \sup_{(f, \mu) \in \mathcal{A}_{\text{safe}, \epsilon}} \Phi_i(f, \mu) \right] \quad (2.7a)$$

$$J_{\text{unsafe}, \epsilon}(\Phi_i) := \left[ \inf_{(f, \mu) \in \mathcal{A}_{\text{unsafe}, \epsilon}} \Phi_i(f, \mu), \sup_{(f, \mu) \in \mathcal{A}_{\text{unsafe}, \epsilon}} \Phi_i(f, \mu) \right]. \quad (2.7b)$$

Note that, in general, these two intervals may be disjoint or may have non-empty intersection; the size of their intersection provides a measure of usefulness of the proposed experiment  $E_i$ . Observe that if experiment  $E_i$  were run, yielding the value  $\Phi_i(G, \mathbb{P})$ ,

then the following conclusions could be drawn:

$$\begin{aligned}
\Phi_i(G, \mathbb{P}) \in J_{\text{safe}, \epsilon}(\Phi_i) \cap J_{\text{unsafe}, \epsilon}(\Phi_i) &\implies \text{no conclusion,} \\
\Phi_i(G, \mathbb{P}) \in J_{\text{safe}, \epsilon}(\Phi_i) \setminus J_{\text{unsafe}, \epsilon}(\Phi_i) &\implies \text{the system is safe,} \\
\Phi_i(G, \mathbb{P}) \in J_{\text{unsafe}, \epsilon}(\Phi_i) \setminus J_{\text{safe}, \epsilon}(\Phi_i) &\implies \text{the system is unsafe,} \\
\Phi_i(G, \mathbb{P}) \notin J_{\text{safe}, \epsilon}(\Phi_i) \cup J_{\text{unsafe}, \epsilon}(\Phi_i) &\implies \text{faulty assumptions,}
\end{aligned}$$

where the last assertion (*faulty assumptions*) means that  $(G, \mathbb{P}) \notin \mathcal{A}$  and follows from the fact that  $\Phi_i(G, \mathbb{P}) \notin J_{\text{safe}, \epsilon}(\Phi_i) \cup J_{\text{unsafe}, \epsilon}(\Phi_i)$  is a contradiction. The validity of the first three assertions is based on the supposition that  $(G, \mathbb{P}) \in \mathcal{A}$ .

In this way, the computational optimization exercise of finding  $J_{\text{safe}, \epsilon}(\Phi_i)$  and  $J_{\text{unsafe}, \epsilon}(\Phi_i)$  for each proposed experiment  $E_i$  provides an objective assessment of which experiments are worth performing: those for which  $J_{\text{safe}, \epsilon}(\Phi_i)$  and  $J_{\text{unsafe}, \epsilon}(\Phi_i)$  are nearly disjoint intervals are worth performing since they are likely to yield conclusive results *vis-à-vis* (de-)certification and conversely, if the intervals  $J_{\text{safe}, \epsilon}(\Phi_i)$  and  $J_{\text{unsafe}, \epsilon}(\Phi_i)$  have a large overlap, then experiment  $E_i$  is not worth performing since it is unlikely to yield conclusive results. Furthermore, the fourth possibility above shows how experiments can rigorously establish that one's assumptions  $\mathcal{A}$  are incorrect. See Figure 2.2 for an illustration.

**Remark 2.3.** For the sake of clarity, we have started this description by defining *experiments* as *functionals*  $\Phi_i$  of  $\mathbb{P}$  and  $G$ . In practice, some experiments may not be functionals of  $\mathbb{P}$  and  $G$  but of related objects. Consider, for instance, the situation where  $(X_1, X_2)$  is a two-dimensional Gaussian vector with zero mean and covariance matrix  $C$ ,  $\mathbb{P}$  is the probability distribution of  $X_1$ , the experiment  $E_2$  determines the variance of  $X_2$  and the information set  $\mathcal{A}$  is  $C \in B$ , where  $B$  is a subset of symmetric positive definite  $2 \times 2$  matrices. The outcome of the experiment  $E_2$  is not a function of the probability distribution  $\mathbb{P}$ ; however, the knowledge of  $\mathbb{P}$  restricts the range of possible outcomes of  $E_2$ . Hence, for some experiments  $E_i$ , the knowledge of  $(G, \mathbb{P})$  does not determine the outcome of the experiment, but only the set of possible outcomes. For those experiments, the description given above can be generalized to situations where  $\Phi_i$  is a *multivalued* functional of  $(G, \mathbb{P})$  determining the set of possible outcomes of the experiment  $E_i$ . This picture can be generalized further by introducing measurement noise, in which case  $(G, \mathbb{P})$  may not determine a deterministic set of possible outcomes, but instead a measure of probability on a set of possible outcomes.

**Example 2.4** (Computational solution of the experimental selection problem). We will now consider again the admissible set  $\mathcal{A}_H$  as given by (1.8). The following example shows that the notion of “best” experiment depends on the admissible safety threshold  $\epsilon$  for  $\mathbb{P}[G \geq a]$ . Suppose that an experiment  $E$  is proposed that will determine the probability mass of the upper half of the velocity range,  $[2.45, 2.8] \text{ km} \cdot \text{s}^{-1}$ ; the corresponding functional  $\Phi$  of study is

$$\Phi(\mu) := \mu[v \geq 2.45 \text{ km} \cdot \text{s}^{-1}],$$

and the proposed experiment  $E$  will determine  $\Phi(\mathbb{P})$  (approximate determinations including measurement and sampling errors can also be handled, but the exact determination



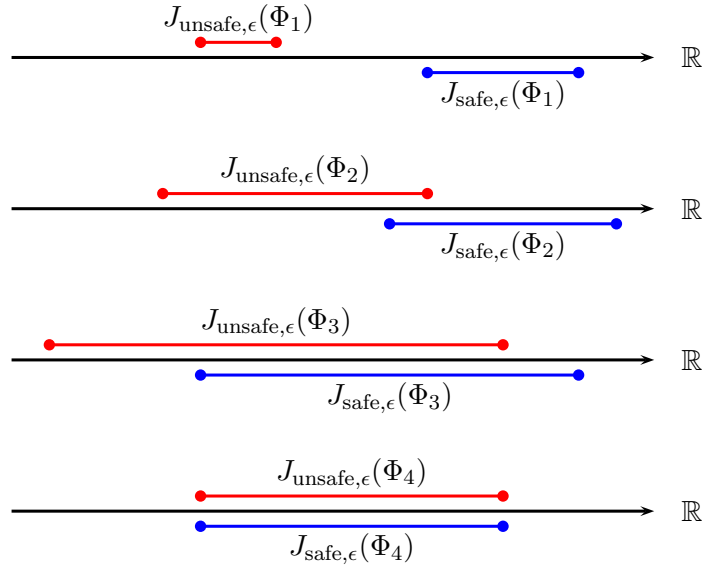


Figure 2.2: A schematic representation of the intervals  $J_{\text{unsafe},\epsilon}(\Phi_i)$  (in red) and  $J_{\text{safe},\epsilon}(\Phi_i)$  (in blue) as defined by (2.7) for four functionals  $\Phi_i$  that might be the subject of an experiment.  $\Phi_1$  is a good candidate for experiment effort, since the intervals do not overlap and hence experimental determination of  $\Phi_1(G, \mathbb{P})$  will certify or de-certify the system;  $\Phi_4$  is not worth investigating, since it cannot distinguish safe scenarios from unsafe ones;  $\Phi_2$  and  $\Phi_3$  are intermediate cases, and  $\Phi_2$  is a better prospect than  $\Phi_3$ .

is done here for simplicity). The corresponding intervals  $J_{\text{safe},\epsilon}(\Phi)$  and  $J_{\text{unsafe},\epsilon}(\Phi)$  as defined by (2.7) and (2.3) are reported in Table 2.3 for various acceptable probabilities of failure  $\epsilon$ . Note that, for larger values of  $\epsilon$ ,  $E$  is a “better” experiment in the sense that it can distinguish safe scenarios from unsafe ones (see also Figure 2.2); for smaller values of  $\epsilon$ ,  $E$  is a poor experiment. In any case, since the intersection  $J_{\text{safe},\epsilon}(\Phi) \cap J_{\text{unsafe},\epsilon}(\Phi)$  is not empty,  $E$  is not an ideal experiment.

It is important to note that the optimization calculations necessary to compute the intervals  $J_{\text{safe},\epsilon}(\Phi)$  and  $J_{\text{unsafe},\epsilon}(\Phi)$  are simplified by the application of Theorem 4.1: in this case, the objective function of  $\mu$  is  $\mu[v \geq 2.45]$  instead of  $\mu[H = 0]$ , but the constraints are once again linear inequalities on generalized moments of the optimization variable  $\mu$ .

**On the number of total evaluations on  $H$ .** Recall that, for simplicity, we have assumed the response function  $G$  to be known and given by  $H$ . A large number of evaluations of  $H$  has been used in Table 2.3 to ensure convergence towards the global optimum. It is important to observe that those evaluations of  $H$  are not (actual, costly) experiments but (cheap) numerical evaluations of equation (1.5). More precisely, the

	$J_{\text{safe},\epsilon}(\Phi)$		$J_{\text{unsafe},\epsilon}(\Phi)$	
	inf	sup	inf	sup
$\epsilon = 0.100$	0.000	1.000	0.000	0.900
iterations until numerical convergence	40	40	40	300
total evaluations of $H$	1,000	1,000	1,000	8,000
$\epsilon = 0.200$	0.000	1.000	0.000	0.800
iterations until numerical convergence	40	40	40	400
total evaluations of $H$	1,000	1,000	1,000	12,000
$\epsilon = 0.300$	0.000	1.000	0.000	0.599
iterations until numerical convergence	40	40	40	1000
total evaluations of $H$	1,000	1,000	1,000	33,000

Table 2.3: The results of the calculation of the intervals  $J_{\text{safe},\epsilon}(\Phi)$  and  $J_{\text{unsafe},\epsilon}(\Phi)$  for the functional  $\Phi(\mu) := \mu[v \geq 2.45 \text{ km} \cdot \text{s}^{-1}]$ . Note that, as the acceptable probability of system failure,  $\epsilon$ , increases, the two intervals overlap less, so experimental determination of  $\Phi(\mathbb{P})$  would be more likely to yield a decisive conservative certification of the system as safe or unsafe; the computational cost of this increased decisiveness is a greater number of function evaluations in the optimization calculations. All computational cost figures are approximate.

method for selecting new best experiments does not require new experiments; i.e., it relies entirely on the information set  $\mathcal{A}$  (which contains the information gathered from previous experiments). Hence those evaluations should not be *viewed as* “information gained from Monte Carlo samples” but as “pure CPU processing time”. In situations where the numerical evaluation of  $H$  is expensive, one can introduce its cost in the optimization loop. An investigation of the best algorithm to perform the numerical optimization with the least number of function evaluations is a worthwhile subject but is beyond the scope of the present paper. Observe also that the method proposed in Section 2 does not rely on the exact knowledge of the response function  $G$ . More precisely, in situations where the response function is unknown, the selection of next best experiments is still entirely computational, and based upon previous data/information gathered on  $G$  enforced as constraints in a numerical optimization algorithm. More precisely, in those situations, the optimization algorithm may require the numerical evaluation of a large number of admissible functions  $f$  (compatible with the prior information available on  $G$ ) but it does not require any new evaluation of  $G$ .

In situations where  $H$  is (numerically) expensive to evaluate, one would have to include the cost of these evaluations in the optimization loop and use fast algorithms exploiting the multi-linear structures associated with the computation of safe and unsafe intervals. Here we have used a genetic algorithm because its robustness. This algorithm typically converges at 10% of the total number of evaluations of  $H$  given in the last row

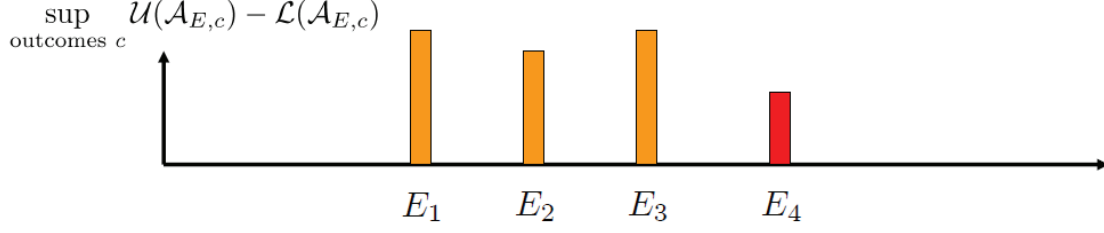


Figure 2.3: A schematic representation of the size of the prediction intervals  $\sup_{\text{outcomes } c} (\mathcal{U}(\mathcal{A}_{E,c}) - \mathcal{L}(\mathcal{A}_{E,c}))$  in the worst case with respect to outcome  $c$ .  $E_4$  is the most predictive experiment.

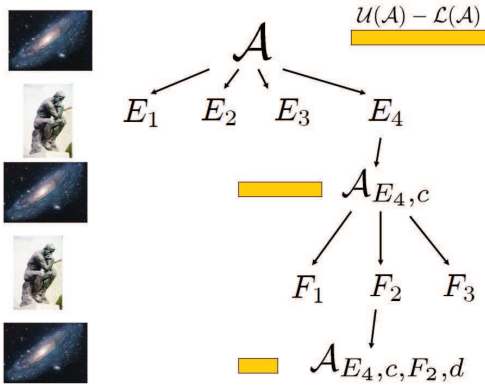
of Table 2.3 but we have increased the number of iterations tenfold to guarantee a robust result. The investigation of efficient optimization algorithms exploiting the multi-linear structures of OUQ optimization problems is of great interest and beyond the immediate scope of this paper.

**Selection of the most predictive experiment.** The computation of safe and unsafe intervals described in the previous paragraph allows of the selection of the most selective experiment. If our objective is to have an “accurate” prediction of  $\mathbb{P}[G(X) \geq a]$ , in the sense that  $\mathcal{U}(\mathcal{A}) - \mathcal{L}(\mathcal{A})$  is small, then one can proceed as follows. Let  $\mathcal{A}_{E,c}$  denote those scenarios in  $\mathcal{A}$  that are compatible with obtaining outcome  $c$  from experiment  $E$ . An experiment  $E^*$  that is most predictive, even in the worst case, is defined by a minmax criterion: we seek (see Figure 2.3)

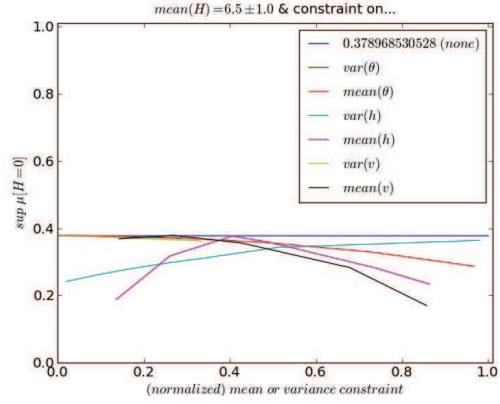
$$E^* \in \arg \min_{\text{experiments } E} \left( \sup_{\text{outcomes } c} (\mathcal{U}(\mathcal{A}_{E,c}) - \mathcal{L}(\mathcal{A}_{E,c})) \right) \quad (2.8)$$

The idea is that, although we can not predict the precise outcome  $c$  of an experiment  $E$ , we can compute a worst-case scenario with respect to  $c$ , and obtain an optimal bound for the minimum decrease in our prediction interval for  $\mathbb{P}[G(X) \geq a]$  based on the (yet unknown) information gained from experiment  $E$ . Again, the theorems given in this paper can be applied to reduce this kind of problem. Finding  $E^*$  is a bigger problem than just calculating  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$ , but the presumption is that computer time is cheaper than experimental effort.

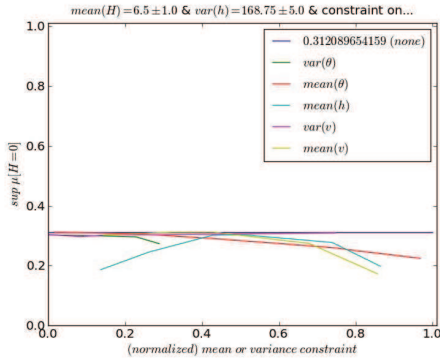
**Planning of campaigns of experiments.** The idea of experimental selection can be extended to plan several experiments in advance, i.e. to plan campaigns of experiments. This aspect can be used to assess the safety or the design of complex systems in a minimal number of experiments (and also to predict bounds on the total number of required experiments). Just as a good chess player thinks several moves ahead, our framework allows for the design of increasingly sophisticated and optimal sequences of experiments that can be performed to measure key system variables. The implementation of this strategy corresponds to a min max game “played against the Universe”



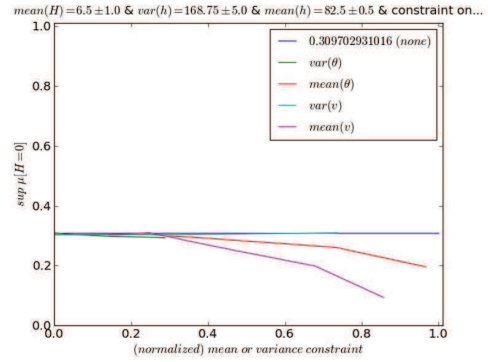
(a) Playing Chess against the universe



(b) Let's play Clue, Round 1



(c) Let's play Clue, Round 2



(d) Let's play Clue, Round 3

Figure 2.4: Subfigure (a): Playing Chess Against the Universe. We choose which experiment  $E$  to perform and the universe selects the outcome  $c$ . Our objective is to minimize  $\mathcal{U}(\mathcal{A}) - \mathcal{L}(\mathcal{A})$ . In the first round our possible moves correspond to a choice between experiments  $E_1, E_2, E_3$  and  $E_4$ . We perform experiment  $E_4$ , the outcome  $c$  of that experiment (selected by the Universe) transforms the admissible into  $\mathcal{A}_{E_4,c}$ . In the second round, our possible moves correspond to a choice between experiments  $F_1, F_2$  and  $F_3$ . As in the game of Chess, several moves can be planned in advance by solving min max optimization problems, and the exponential increase of the number branches of the game tree can be kept under control by exploring only a subset of (best) moves. Subfigures (b), (c) and (d): Let's play Clue.

(Subfigure 2.4(a)). The well-known games of Clue/Cluedo and Twenty Questions are better analogies than chess for this kind of information game. In that sense, the planning of campaigns of experiments is an infinite-dimensional Clue, played on spaces of admissible scenarios, against our lack of perfect information about reality, and made tractable by the reduction theorems. This aspect calls for more investigation since it has the potential to provide a new approach to the current scientific investigation paradigm, which is based on intuition, expert judgment, and guessing.

**Example 2.5** (Let's play Clue.). In Subfigures 2.4(b), 2.4(c) and 2.4(d) we consider again the admissible set  $\mathcal{A}_H$  as given by (1.8) and select three most predictive experiments, sequentially, choosing the second one after having observed the outcome of the first one. The list of possible experiments corresponds to measuring the mean or variance of thickness  $h$ , obliquity  $\theta$  or velocity  $v$ . Subfigures 2.4(b), 2.4(c) and 2.4(d) show  $\mathcal{U}(\mathcal{A}_{H,E,c})$  for each of these experiments as a function of the re-normalized outcome value  $c$ . Since, in this example, we always have  $\mathcal{L}(\mathcal{A}_{H,E,c}) = 0$ ,  $\mathcal{U}(\mathcal{A}_{H,E,c})$  corresponds to the size of the prediction interval for the probability of non-perforation given the information that the outcome of experiment  $E$  is  $c$ . Given the results shown in Subfigure 2.4(b) we select to measure the variance of thickness as our first best experiment. Note that this selection does not require the knowledge of what the experiment *will* yield but only the knowledge of what the experiment *can* yield: we identify as the optimal next experiment the one that is most informative in the minimax sense, i.e. of all its possible outcomes (this does not require the knowledge of what the experiment will yield), its least informative outcome is more informative than the least informative outcome of any other candidate experiment. Although not necessary, this selection can, possibly, be guided by a model of reality (i.e. in this case a model for the probability distributions of  $h, \theta, v$ ). Used in this manner, an accurate model will reduce the number of experiments required for certification and an inaccurate model will lead to a relatively greater number of experiments (but not to erroneous bounds). Subfigure 2.4(c) is based on the information contained in  $\mathcal{A}_H$  and bounds on the variance of thickness (obtained from the first experiment). Our selection as second experiment is to measure the mean of thickness (leading to Subfigure 2.4(d)).

### 3 Generalizations and Comparisons

#### 3.1 Prediction, extrapolation, verification and validation

In the previous section, the OUQ framework was described as it applies to the the certification problem (1.1). We will now show that many important UQ problems, such as prediction, verification and validation, can be formulated as certification problems. This is similar to the point of view of [5], in which formulations of many problem objectives in reliability are shown to be representable in a unified framework.

A **prediction problem** can be formulated as, given  $\epsilon$  and (possibly incomplete) information on  $\mathbb{P}$  and  $G$ , finding a smallest  $b - a$  such that

$$\mathbb{P}[a \leq G(X) \leq b] \geq 1 - \epsilon, \quad (3.1)$$

which, given the admissible set  $\mathcal{A}$ , is equivalent to solving

$$\inf \left\{ b - a \mid \inf_{(f, \mu) \in \mathcal{A}} \mu[a \leq f(X) \leq b] \geq 1 - \epsilon \right\}. \quad (3.2)$$

Observe that  $[a, b]$  can be interpreted as an optimal interval of confidence for  $G(X)$  (although  $b - a$  is minimal,  $[a, b]$  may not be unique), in particular, with probability at least  $1 - \epsilon$ ,  $G(X) \in [a, b]$ .

In many applications the regime where experimental data can be taken is different than the deployment regime where prediction or certification is sought, and this is commonly referred to as the **extrapolation problem**. For example, in materials modeling, experimental tests are performed on materials, and the model run for comparison, but the desire is that these results tell us something where experimental tests are impossible, or extremely expensive to obtain.

In most applications, the response function  $G$  may be approximated via a (possibly numerical) model  $F$ . Information on the relation between the model  $F$  and the response function  $G$  that it is designed to represent (i.e. information on  $(x, F(x), G(x))$ ) can be used to restrict (constrain) the set  $\mathcal{A}$  of admissible scenarios  $(G, \mathbb{P})$ . This information may take the form of a bound on some distance between  $F$  and  $G$  or a bound on some complex functional of  $F$  and  $G$  [52, 78]. Observe that, in the context of the certification problem (1.1), the value of the model can be measured by changes induced on the optimal bounds  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$ . The problem of quantifying the relation (possibly the distance) between  $F$  and  $G$  is commonly referred to as the **validation problem**. In some situations  $F$  may be a numerical model involving millions of lines of code and (possibly) space-time discretization. The quantification of the uncertainty associated with the possible presence of bugs and discretization approximations is commonly referred to as the **verification problem**. Both, the validation and the verification problem, can be addressed in the OUQ framework by introducing information sets describing relations between  $G$ ,  $F$  and the code.

### 3.2 Comparisons with other UQ methods

We will now compare OUQ with other widely used UQ methods and consider the certification problem (1.1) to be specific.

- Assume that  $n$  independent samples  $Y_1, \dots, Y_n$  of the random variable  $G(X)$  are available (i.e.  $n$  independent observations of the random variable  $G(X)$ , all distributed according to the measure of probability  $\mathbb{P}$ ). If  $\mathbb{1}[Y_i \geq a]$  denotes the random variable equal to one if  $Y_i \geq a$  and equal to zero otherwise, then

$$p_n := \frac{\sum_{i=1}^n \mathbb{1}[Y_i \geq a]}{n} \quad (3.3)$$

is an unbiased estimator of  $\mathbb{P}[G(X) \geq a]$ . Furthermore, as a result of Hoeffding's concentration inequality [33], the probability that  $p_n$  deviates from  $\mathbb{P}[G(X) \geq a]$

(its mean) by at least  $\epsilon/2$  is bounded from above by  $\exp(-\frac{n}{2}\epsilon^2)$ . It follows that if the number of samples  $n$  is large enough (of the order of  $\frac{1}{\epsilon^2} \log \frac{1}{\epsilon}$ ), then the certification of (1.1) can be obtained through a Monte Carlo estimate (using  $p_n$ ). As this example shows, **Monte Carlo strategies** [50] are simple to implement and do not necessitate prior information on the response function  $G$  and the measure  $\mathbb{P}$  (other than the i.i.d. samples). However, they require a large number of (independent) samples of  $G(X)$  which is a severe limitation for the certification of rare events (the  $\epsilon = 10^{-9}$  of the aviation industry would [83, 15] necessitate  $O(10^{18})$  samples). Additional information on  $G$  and  $\mathbb{P}$  can, in principle, be included (in a limited fashion) in Monte Carlo strategies via importance and weighted sampling [50] to reduce the number of required samples.

- The number of required samples can also be reduced to  $\frac{1}{\epsilon}(\ln \frac{1}{\epsilon})^d$  using **Quasi-Monte Carlo Methods**. We refer in particular to the Koksma–Hlawka inequality [64], to [82] for multiple integration based on lattice rules and to [81] for a recent review. We observe that these methods require some regularity (differentiability) condition on the response function  $G$  and the possibility of sampling  $G$  at pre-determined points  $X$ . Furthermore, the number of required samples blows-up at an exponential rate with the dimension  $d$  of the input vector  $X$ .
- If  $G$  is regular enough and can be sampled at at pre-determined points, and if  $X$  has a known distribution, then **stochastic expansion methods** [30, 29, 101, 4, 26, 19] can reduce the number of required samples even further (depending on the regularity of  $G$ ) provided that the dimension of  $X$  is not too high [94, 14]. However, in most applications, only incomplete information on  $\mathbb{P}$  and  $G$  is available and the number of available samples on  $G$  is small or zero.  $X$  may be of high dimension, and may include uncontrollable variables and unknown unknowns (unknown input parameters of the response function  $G$ ).  $G$  may not be the solution of a PDE and may involve interactions between singular and complex processes such as (for instance) dislocation, fragmentation, phase transitions, physical phenomena in untested regimes, and even human decisions. We observe that in many applications of Stochastic Expansion methods  $G$  and  $\mathbb{P}$  are assumed to be perfectly known and UQ reduces to computing the push forward of the measure  $\mathbb{P}$  via the response (transfer) function  $I_{\geq a} \circ G$  (to a measure on two points, in those situations  $\mathcal{L}(\mathcal{A}) = \mathbb{P}[G \geq a] = \mathcal{U}(\mathcal{A})$ ).
- The investigation of variations of the response function  $G$  under variations of the input parameters  $X_i$ , commonly referred to as **sensitivity analysis** [75, 76], allows for the identification of critical input parameters. Although helpful in estimating the robustness of conclusions made based on specific assumptions on input parameters, sensitivity analysis, in its most general form, has not been targeted at obtaining rigorous upper bounds on probabilities of failures associated with certification problems (1.1). However, single parameter oscillations of the function  $G$  (as defined by (1.10)) can be seen as a form of non-linear sensitivity analysis leading to bounds on  $\mathbb{P}[G \geq a]$  via McDiarmid’s concentration inequality [57, 58].



These bounds can be made sharp by partitioning the input parameter space along maximum oscillation directions and computing sub-diameters on sub-domains [91].

- If  $\mathcal{A}$  is expressed probabilistically through a prior (an a priori measure of probability) on the set possible scenarios  $(f, \mu)$  then **Bayesian inference** [48, 7] could in principle be used to estimate  $\mathbb{P}[G \geq a]$  using the posterior measure of probability on  $(f, \mu)$ . This combination between OUQ and Bayesian methods avoids the necessity to solve the possibly large optimization problems (2.4) and it also greatly simplifies the incorporation of sampled data thanks to the Bayes rule. However, oftentimes, priors are not available or their choice involves some degree of arbitrariness that is incompatible with the certification of rare events. Priors may become asymptotically irrelevant (in the limit of large data sets) but, for small  $\epsilon$ , the number of required samples can be of the same order as the number required by Monte-Carlo methods [79].

When unknown parameters are estimated using priors and sampled data, it is important to observe that the convergence of the Bayesian method may fail if the underlying probability mechanism allows an infinite number of possible outcomes (e.g., estimation of an unknown probability on  $\mathbb{N}$ , the set of all natural numbers) [17]. In fact, in these infinite-dimensional situations, this lack of convergence (commonly referred to as inconsistency) is the rule rather than the exception [18]. As emphasized in [17], *as more data comes in, some Bayesian statisticians will become more and more convinced of the wrong answer*.

We also observe that, for complex systems, the computation of posterior probabilities has been made possible thanks to advances in computer science. We refer to [89] for a (recent) general (Gaussian) framework for Bayesian inverse problems and [6] for a rigorous UQ framework based on probability logic with Bayesian updating. Just as Bayesian methods would have been considered computationally infeasible 50 years ago but are now common practice, OUQ methods are now becoming feasible and will only increase in feasibility with the passage of time and advances in computing.

- The combination of structural optimization (in various fields of engineering) to produce an optimal design given the (deterministic) worst-case scenario has been referred to as **Optimization and Anti-Optimization** [27] (we also refer to **critical excitation** in seismic engineering [21]). The main difference between OUQ and Anti-optimization lies in the fact that the former is based on an optimization over (admissible) functions and measures  $(f, \mu)$ , while the latter only involves an optimization over  $f$ . Because of its robustness, many engineers have adopted the (deterministic) worst-case scenario approach to UQ (we refer to chapter 10 of [27]) when a high reliability is required. It is noted in [27] that the reason why *probabilistic methods do not find appreciation among theoreticians and practitioners alike* lies in the fact that “probabilistic reliability studies involve assumptions on the probability densities, whose knowledge regarding relevant input quantities is central”. It is also observed that strong assumptions on  $\mathbb{P}$  may lead to GIGO



(“garbage in–garbage out”) situations for small certification thresholds  $\epsilon$  when reliability estimates and probabilities of failure are very sensitive to small deviations in probability densities. On the other hand, UQ methods based on deterministic worst-case scenarios are oftentimes “too pessimistic to be practical” [21, 27]. We suggest that by allowing for very weak assumptions on probability measures, OUQ methods can lead to bounds on probabilities of failure that are both reliable and practical. Indeed, when applied to complex systems involving a large number of variables, deterministic worst-case methods do not take into account the *improbability* that these (possibly independent or weakly correlated) variables conspire to produce a failure event.

The certification problem (1.1) exhibits one of the main difficulties that face UQ practitioners: many theoretical methods are available, but they require assumptions or conditions that, oftentimes, are not satisfied by the application. More precisely, the characteristic elements distinguishing these different methods are the assumptions upon which they are based, and some methods will be more efficient than others depending on the validity of those assumptions. UQ applications are also characterized by a set of assumptions/information on the response function  $G$  and measure  $\mathbb{P}$ , which varies from application to application. Hence, on the one hand, we have a list of theoretical methods that are applicable or efficient under very specific assumptions; on the other hand, most applications are characterized by an information set or assumptions that, in general, do not match those required by these theoretical methods. It is hence natural to pursue the development of a rigorous framework that does not add inappropriate assumptions or discard information.

We also observe that the effectiveness of different UQ methods cannot be compared without reference to the available information (some methods will be more efficient than others depending on those assumptions). For the hypervelocity impact example of Subsection 1.2, none of the methods mentioned above can be used without adding (arbitrary) assumptions on probability densities or discarding information on the mean value or independence of the input parameters. We also observe that it is by placing information at the center of UQ that the proposed OUQ framework allows for the identification of best experiments. Without focus on the available information, UQ methods are faced with the risk of propagating inappropriate assumptions and producing a sophisticated answer to the wrong question. These distortions of the information set may be of limited impact on certification of common events but they are also of critical importance for the certification of rare events.

### 3.3 OUQ with random sample data

For the sake of clarity, we have started the description of OUQ with deterministic information and assumptions (i.e. when  $\mathcal{A}$  is a deterministic set of functions and probability measures). In many applications, however, some of the information arrives in the form of random samples. The addition of such sample data to the available information and assumptions leads to non-trivial theoretical questions that are of practical impor-

tance beyond their fundamental connections with information theory and nonparametric statistics. In particular, while the notion of an optimal bound (2.4) is transparent and unambiguous, the notion of an optimal bound on  $\mathbb{P}[G(X) \geq a]$  in presence of sample data is not immediately obvious and should be treated with care. This is a very delicate topic, a full treatment of which we shall defer to a future work.

## 4 Reduction of OUQ Optimization Problems

In general, the lower and upper values

$$\begin{aligned}\mathcal{L}(\mathcal{A}) &:= \inf_{(f,\mu) \in \mathcal{A}} \mu[f(X) \geq a] \\ \mathcal{U}(\mathcal{A}) &:= \sup_{(f,\mu) \in \mathcal{A}} \mu[f(X) \geq a]\end{aligned}$$

are each defined by a non-convex and infinite-dimensional optimization problem, the solution of which poses significant computational challenges. These optimization problems can be considered to be a generalization of Chebyshev inequalities. The history of the classical inequalities can be found in [42], and some generalizations in [13] and [98]; in the latter works, the connection between Chebyshev inequalities and optimization theory is developed based on the work of Mulholland and Rogers [62], Godwin [32], Isii [37, 38, 39], Olkin and Pratt [65], Marshall and Olkin [54], and the classical Markov–Krein Theorem [42, pages 82 & 157], among others. The Chebyshev-type inequalities defined by  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$  are a further generalization to independence assumptions, more general domains, more general systems of moments, and the inclusion of classes of functions, in addition to the probability measures, in the optimization problem. Moreover, although our goal is the computation of these values, and not an analytic expression for them, the study of probability inequalities should be useful in the reduction and approximation of these values. Without providing a survey of this large body of work, we mention the field of majorization, as discussed in Marshall and Olkin [55], the inequalities of Anderson [3], Hoeffding [33], Joe [40], Bentkus *et al.* [11], Bentkus [9, 10], Pinelis [69, 70], and Boucheron, Lugosi and Massart [16]. Moreover, the solution of the resulting nonconvex optimization problems should benefit from duality theories for nonconvex optimization problems such as Rockafellar [74] and the development of convex envelopes for them, as can be found, for example, in Rikun [73] and Sherali [80]. Finally, since Pardalos and Vavasis [67] show that quadratic programming with one negative eigenvalue is NP-hard, we expect that some OUQ problems may be difficult to solve.

Let us now return to the earlier simple example of an admissible set  $\mathcal{A}_1$  in (2.1): the (non-unique) extremizers of the OUQ problem with the admissible set  $\mathcal{A}_1$  all have the property that the support of the push-forward measure  $f_*\mu$  on  $\mathbb{R}$  contains at most two points, i.e.  $f_*\mu$  is a convex combination of at most two Dirac delta measures (we recall that  $\#\text{supp}(f_*\mu)$  is the number of points forming the support of  $f_*\mu$ ):

$$\sup_{(f,\mu) \in \mathcal{A}_1} \mu[f(X) \geq a] = \sup_{\substack{(f,\mu) \in \mathcal{A}_1 \\ \#\text{supp}(f_*\mu) \leq 2}} \mu[f(X) \geq a].$$

The optimization problem on the left-hand side is an infinite-dimensional one, whereas the optimization problem on the right-hand side is amenable to finite-dimensional parametrization for each  $f$ . Furthermore, for each  $f$ , only the two values of  $f$  at the support points of the two Dirac measures are relevant to the problem. The aim of this section is to show that a large class of OUQ problems — those governed by independence and linear inequality constraints on the moments, — are amenable to a similar finite-dimensional reduction, and that a priori upper bounds can be given on the number of Dirac delta masses that the reduction requires.

To begin with, we first show that an important class of optimization problems over the space of  $m$ -fold product measures can be reduced to optimization over products of finite convex combinations of Dirac masses ( $m$  is the number of random input variables). Consequently, we then show in Corollary 4.4 that OUQ optimization problems where the admissible set is defined as a subset of function-measure pairs  $(f, \mu)$  that satisfy generalized moment constraints  $\mathbf{G}_f(\mu) \leq 0$  can also be reduced from the space of measures to the products of finite convex combinations of Dirac masses. Theorem 4.7 shows that, when all the constraints are generalized moments of functions of  $f$ , the search space  $\mathcal{G}$  of functions can be further reduced to a search over functions on an  $m$ -fold product of finite discrete spaces, and the search over  $m$ -fold products of finite convex combinations of Dirac masses can be reduced to a search over the products of probability measures on this  $m$ -fold product of finite discrete spaces. This latter reduction completely eliminates dependency on the coordinate positions in  $\mathcal{X}$ . Theorem 4.7 is then used in Proposition 4.8 to obtain an optimal McDiarmid inequality through the formulation of an appropriate OUQ optimization problem followed by the above-mentioned reductions to an optimization problem on the product of functions on  $\{0, 1\}^m$  with the  $m$ -fold products of measures on  $\{0, 1\}^m$ . This problem is then further reduced, by Theorem 4.9, to an optimization problem on the product of the space of subsets (power set) of  $\{0, 1\}^m$  with the product measures on  $\{0, 1\}^m$ . Finally, we obtain analytic solutions to this last problem for  $m = 1, 2, 3$ , thereby obtaining an optimal McDiarmid inequality in these cases. We also obtain an asymptotic formula for general  $m$ . Moreover, the solution for  $m = 2$  indicates important information regarding the diameter parameters  $D_1$  and  $D_2$  (we refer to Example 2.2). For example, if  $D_2$  is sufficiently smaller than  $D_1$ , then the optimal bound only depends on  $D_1$  and therefore, any decrease in  $D_2$  does not improve the inequality. See Subsection 10.1 for the proofs of the results in this section.

#### 4.1 Reduction of OUQ

For a topological space  $\mathcal{X}$ , let  $\mathcal{F}_{\mathcal{X}}$  (or simply  $\mathcal{F}$ ) denote the space of real-valued (Borel) measurable functions on  $\mathcal{X}$ , and let  $\mathcal{M}(\mathcal{X})$  denote the set of Borel probability measures on  $\mathcal{X}$ . Denote the process of integration with respect to a measure  $\mu$  by  $\mathbb{E}_{\mu}$ , and let

$$\Delta_k(\mathcal{X}) := \left\{ \sum_{j=0}^k \alpha_j \delta_{x^j} \left| x^j \in \mathcal{X}, \alpha_j \geq 0 \text{ for } j = 0, \dots, k \text{ and } \sum_{j=0}^k \alpha_j = 1 \right. \right\}$$

denote the set of  $(k+1)$ -fold convex combinations of Dirac masses. When  $\mathcal{X} = \prod_{i=1}^m \mathcal{X}_i$  is a product of topological spaces, and we speak of measurable functions on the product  $\mathcal{X}$ , we mean measurable with respect to the product  $\sigma$ -algebra and not the Borel  $\sigma$ -algebra of the product. For more discussion of this delicate topic, see e.g. [41]. The linear equality and inequality constraints on our optimization problems will be encoded in the following measurable functions:

$$g'_j: \mathcal{X} \rightarrow \mathbb{R} \text{ for } j = 1, \dots, n',$$

and, for each  $i = 1, \dots, m$ ,

$$g_j^i: \mathcal{X}_i \rightarrow \mathbb{R} \text{ for } j = 1, \dots, n_i.$$

Let  $\mathcal{M}^{\mathbf{G}} \subseteq \mathcal{M}^m(\mathcal{X})$  denote the set of products of Borel measures for which these all these functions are integrable with finite integrals. We use the compact notation  $\mathbf{G}(\mu) \leq 0$  to indicate that  $\mu \in \mathcal{M}^{\mathbf{G}}$  and that

$$\begin{aligned} \mathbb{E}_\mu[g'_j] &\leq 0, \quad \text{for all } j = 1, \dots, n', \\ \mathbb{E}_\mu[g_j^i] &\leq 0, \quad j = 1, \dots, n_i, \text{ for all } i = 1, \dots, m. \end{aligned}$$

Moreover, let  $r: \mathcal{X} \rightarrow \mathbb{R}$  be integrable for all  $\mu \in \mathcal{M}^{\mathbf{G}}$  (possibly with values  $+\infty$  or  $-\infty$ ). For any set  $\mathcal{M} \subseteq \mathcal{M}^{\mathbf{G}}$ , let

$$\mathcal{U}(\mathcal{M}) := \sup_{\mu \in \mathcal{M}} \mathbb{E}_\mu[r],$$

with the convention that the supremum of the empty set is  $-\infty$ .

For a measurable function  $f$ , the map  $\mu \mapsto \mathbb{E}_\mu[f]$  may not be defined, since  $f$  may not be absolutely integrable with respect to  $\mu$ . If it is defined, then it is continuous in the strong topology on measures; however, this topology is too strong to provide any compactness. Moreover, although [2, Theorem 14.5] shows that if  $f$  is a bounded upper semi-continuous function on a metric space, then integration is upper semi-continuous in the weak-\* topology, we consider the case in which  $\mathcal{X}$  may not be metric or compact, and the functions  $f$  may be unbounded and lack continuity properties. The following results heavily use results of Winkler [100] — which follow from an extension of Choquet Theory (see e.g. [68]) by von Weizsäcker and Winkler [99, Corollary 3] to sets of probability measures with generalized moment constraints — and a result of Kendall [45] characterizing cones, which are lattice cones in their own order. These results generalize a result of Karr [43] that requires  $\mathcal{X}$  to be compact, the constraint functions be bounded and continuous, and the constraints to be equalities. The results that follow are remarkable in that they make extremely weak assumptions on  $\mathcal{X}$  and no assumptions on the functions  $f$ . Recall that a Suslin space is the continuous image of a Polish space.

**Theorem 4.1.** *Let  $\mathcal{X} = \prod_{i=1}^m \mathcal{X}_i$  be a product of Suslin spaces and let*

$$\mathcal{M}^m(\mathcal{X}) := \bigotimes_{i=1}^m \mathcal{M}(\mathcal{X}_i)$$

denote the set of products of Borel probability measures on the spaces  $\mathcal{X}_i$ . As above, consider the generalized moment functions  $\mathbf{G}$  and the corresponding finite moment set  $\mathcal{M}^{\mathbf{G}}$ . Suppose that  $r: \mathcal{X} \rightarrow \mathbb{R}$  is integrable for all  $\mu \in \mathcal{M}^{\mathbf{G}}$  (possibly with values  $+\infty$  or  $-\infty$ ). Define the reduced admissible set

$$\mathcal{M}_{\Delta} := \left\{ \mu \in \bigotimes_{i=1}^m \Delta_{n_i+n'}(\mathcal{X}_i) \mid \mathbf{G}(\mu) \leq 0 \right\}.$$

Then, it holds that

$$\mathcal{U}(\mathcal{M}^{\mathbf{G}}) = \mathcal{U}(\mathcal{M}_{\Delta}).$$

Theorem 4.1 says that, on a product  $\mathcal{X}$  of very general spaces  $\mathcal{X}_i$ , optimization problems constrained by  $n'$  linear moment constraints on  $\mathcal{X}$  and  $n_i$  linear moment constraints on each factor space  $\mathcal{X}_i$  achieve their optima among those product measures whose  $i^{\text{th}}$  marginal has support on at most  $n' + n_i + 1$  points of  $\mathcal{X}_i$ .

**Remark 4.2.** Using [99, Corollary 3], this theorem and its consequences below easily generalize from the situation where  $\mathbb{E}_{\mu}[g_k] \leq 0$  for each  $k$  to that in which  $\mathbb{E}_{\mu}[g_k] \in I_k$  for each  $k$ , where  $k$  indexes the constraint functions, and where each  $I_k$  is a closed interval. Consequently, such pairs of linear constraints introduce a requirement for only one Dirac mass, not the two masses that one might expect. Moreover, observe that the condition that the function  $r$  is integrable (possibly with values  $+\infty$  or  $-\infty$ ) for all  $\mu \in \mathcal{M}^{\mathbf{G}}$  is satisfied if  $r$  is non-negative. In particular, this holds when  $r$  is an indicator function of a set, which is our main application in this paper.

**Remark 4.3.** Theorem 4.1 and its consequences below can be expressed more generally in terms of extreme points of sets of measures, whereas in the above case, the extreme points are the Dirac masses. To that end, Dynkin [24] describes more general sets of measures and their extreme points, which can be useful in applications. In particular, one could consider

1. sets of measures that are invariant under a transformation (the extreme points are the ergodic measures);
2. symmetric measures on an infinite product space (the extreme points are the simple product measures);
3. the set of stationary distributions for a given Markov transition function;
4. the set of all Markov processes with a given transition function.

We now apply Theorem 4.1 to obtain the same type of reduction for an admissible set  $\mathcal{A} \subseteq \mathcal{F} \times \mathcal{M}^m(\mathcal{X})$  consisting of pairs of functions and product measures — this is the case for the OUQ optimization problems  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$ . Let  $\mathcal{G} \subseteq \mathcal{F}$  denote a subset of real-valued measurable functions on  $\mathcal{X}$  and consider an admissible set  $\mathcal{A} \subseteq \mathcal{G} \times \mathcal{M}^m(\mathcal{X})$  defined in the following way. For each  $f \in \mathcal{G}$ , let  $\mathbf{G}(f, \cdot)$  denote a family of constraints as in Theorem 4.1 and Remark 4.2. For each  $f \in \mathcal{G}$ , let  $\mathcal{M}^{\mathbf{G}_f} \subseteq \mathcal{M}^m(\mathcal{X})$  denote those product probability measures  $\mu$  such that the moments  $\mathbf{G}(f, \mu)$  are well-defined.

Moreover, for each  $f \in \mathcal{G}$ , let  $r_f: \mathcal{X} \rightarrow \mathbb{R}$  be integrable for all  $\mu \in \mathcal{M}^{\mathbf{G}_f}$  (possibly with values  $+\infty$  or  $-\infty$ ). Define the admissible set

$$\mathcal{A} := \{(f, \mu) \in \mathcal{G} \times \mathcal{M}^m(\mathcal{X}) \mid \mathbf{G}(f, \mu) \leq 0\} \quad (4.1)$$

and define the OUQ optimization problem to be

$$\mathcal{U}(\mathcal{A}) := \sup_{(f, \mu) \in \mathcal{A}} \mathbb{E}_\mu[r_f]. \quad (4.2)$$

**Corollary 4.4.** *Consider the OUQ optimization problem (4.2) and define the reduced admissible set  $\mathcal{A}_\Delta \subseteq \mathcal{A}$  by*

$$\mathcal{A}_\Delta := \left\{ (f, \mu) \in \mathcal{G} \times \bigotimes_{i=1}^m \Delta_{n_i+n'}(\mathcal{X}_i) \mid \mathbf{G}(f, \mu) \leq 0 \right\}. \quad (4.3)$$

Then, it holds that

$$\mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_\Delta).$$

**Remark 4.5.** Corollary 4.4 is easily generalized to the case where for each  $f \in \mathcal{G}$ ,  $i$ , and fixed  $\mu_j, j \neq i$ ,  $\mathbf{G}(f, \mu_1, \dots, \mu_i, \dots, \mu_m)$  has affine dimension at most  $m_i$  as  $\mu_i$  varies. In this case

$$\mathcal{A}_\Delta := \left\{ (f, \mu) \in \mathcal{G} \times \bigotimes_{i=1}^m \Delta_{m_i}(\mathcal{X}_i) \mid \mathbf{G}(f, \mu) \leq 0 \right\}.$$

**Remark 4.6.** Linear moment constraints on the factor spaces  $\mathcal{X}_i$  allow to consider information sets with independent random variables  $X_1, \dots, X_m$  and weak constraints on the probability measure of the variables  $X_i$ . An example of such an admissible set is the one associated with Bernstein inequalities [12], in which a priori bounds are given on the variances of the variables  $X_i$ .

## 4.2 Generalized moments of the response function

We now consider the case where the function  $r_f := r \circ f$  is defined through composition with a measurable function  $r$ , and all  $n$  constraints are determined by compositions  $g'_j := g_j \circ f$ , with  $j = 1, \dots, n$ , of the function  $f$ . Hence, the symbol  $\mathbf{G}(f, \mu)$  will mean that all functions  $g_j \circ f$  are  $\mu$  integrable and will represent the values  $\mathbb{E}_\mu[g_j \circ f]$  for  $j = 1, \dots, n$ . That is, we have the admissible set

$$\mathcal{A} := \{(f, \mu) \in \mathcal{G} \times \mathcal{M}^m(\mathcal{X}) \mid \mathbf{G}(f, \mu) \leq 0\} \quad (4.4)$$

and the optimization problem

$$\mathcal{U}(\mathcal{A}) := \sup_{(f, \mu) \in \mathcal{A}} \mathbb{E}_\mu[r \circ f] \quad (4.5)$$

as in (4.2). However, in this case, the fact that the criterion function  $r \circ f$  and the constraint functions  $g_j \circ f$  are compositions of the function  $f$  permits a finite-dimensional

reduction of the space of functions  $\mathcal{G}$  to a space of functions on  $\{0, \dots, n\}^m$  and a reduction of the space of  $m$ -fold products of finite convex combinations of Dirac masses to the space of product measures on  $\{0, \dots, n\}^m$ . This reduction completely eliminates dependency on the coordinate positions in  $\mathcal{X}$ .

Formulating this result precisely will require some additional notation. By the Well-Ordering Theorem, there exists a well-ordering of each  $\mathcal{X}_i$ . Suppose that a total ordering of the elements of the spaces  $\mathcal{X}_i$  for  $i = 1, \dots, m$  is specified. Let  $\mathcal{N} := \{0, \dots, n\}$  and  $\mathcal{D} := \{0, \dots, n\}^m = \mathcal{N}^m$ . Every element  $\mu \in \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i)$  is a product  $\mu = \bigotimes_{i=1}^m \mu_i$  where each factor  $\mu_i$  is a convex sum of  $n + 1$  Dirac masses indexed according to the ordering; that is,

$$\mu_i = \sum_{k=0}^n \alpha_k^i \delta_{x_i^k}$$

for some  $\alpha_1^i, \dots, \alpha_n^i \geq 0$  with unit sum and some  $x_i^1, \dots, x_i^n \in \mathcal{X}_i$  such that, with respect to the given ordering of  $\mathcal{X}_i$ ,

$$x_i^1 \leq x_i^2 \leq \dots \leq x_i^n.$$

Let  $\mathcal{F}_{\mathcal{D}}$  denote the real linear space of real functions on  $\mathcal{D} = \{0, \dots, n\}^m$  and consider the mapping

$$\mathbb{F}: \mathcal{F} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i) \rightarrow \mathcal{F}_{\mathcal{D}}$$

defined by

$$(\mathbb{F}(f, \mu))(i_1, i_2, \dots, i_m) = f(x_1^{i_1}, x_2^{i_2}, \dots, x_m^{i_m}), \quad i_k \in \mathcal{N}, k = 1, \dots, m.$$

$\mathbb{F}$  represents the values of the function  $f$  at the Dirac masses in  $\mu$ , but does not carry information regarding the positions of the Dirac masses or their weights.

**Theorem 4.7.** *Consider the admissible set  $\mathcal{A}$  and optimization problem  $\mathcal{U}(\mathcal{A})$  defined in (4.4) and (4.5) where  $r \circ f$  is integrable (possibly with values  $+\infty$  or  $-\infty$ ) for all product measures. For a subset  $\mathcal{G}_{\mathcal{D}} \subseteq \mathcal{F}_{\mathcal{D}}$ , define the admissible set*

$$\mathcal{A}_{\mathcal{D}} = \{(h, \alpha) \in \mathcal{G}_{\mathcal{D}} \times \mathcal{M}^m(\mathcal{D}) \mid \mathbb{E}_{\alpha}[g_i \circ h] \leq 0 \text{ for all } j = 1, \dots, n\} \quad (4.6)$$

and the optimization problem

$$\mathcal{U}(\mathcal{A}_{\mathcal{D}}) := \sup_{(h, \alpha) \in \mathcal{A}_{\mathcal{D}}} \mathbb{E}_{\alpha}[r \circ h].$$

If

$$\mathbb{F} \left( \mathcal{G} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i) \right) = \mathcal{G}_{\mathcal{D}},$$

then it holds that

$$\mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_{\mathcal{D}}).$$

When the constraint set also includes functions which are not compositions with  $f$ , then Theorem 4.7 does not apply. Although it does appear that results similar to Theorem 4.7 can be obtained, we leave that as a topic for future work.

### 4.3 Application to McDiarmid's inequality

Theorem 4.7 can be applied to the situation of McDiarmid's inequality in order to obtain an optimal solution for that problem. Let  $D_i \geq 0$  for  $i = 1, \dots, m$  and define

$$\mathcal{G} := \{f \in \mathcal{F} \mid \text{Osc}_i(f) \leq D_i \text{ for each } i = 1, \dots, m\}, \quad (4.7)$$

where

$$\text{Osc}_i(f) := \sup_{(x_1, \dots, x_m) \in \mathcal{X}} \sup_{x'_i \in \mathcal{X}_i} |f(\dots, x_i, \dots) - f(\dots, x'_i, \dots)|.$$

We have a product probability measure  $\mathbb{P}$  on  $\mathcal{X}$  and a measurable function  $H: \mathcal{X} \rightarrow \mathbb{R}$  such that  $H \in \mathcal{G}$ . Suppose that we have an upper bound

$$\mathbb{P}[H - \mathbb{E}_{\mathbb{P}}[H] \geq a] \leq \mathbb{H}(a, \mathcal{G}) \text{ for all } H \in \mathcal{G}. \quad (4.8)$$

It follows that if  $H \in \mathcal{G}$  and  $\mathbb{E}_{\mathbb{P}}[H] \leq 0$ , then

$$\mathbb{P}[H \geq a] \leq \mathbb{P}[H - \mathbb{E}_{\mathbb{P}}[H] \geq a] \leq \mathbb{H}(a, \mathcal{G}) \text{ for all } H \in \mathcal{G} \text{ with } \mathbb{E}_{\mathbb{P}}[H] \leq 0.$$

On the other hand, suppose that

$$\mathbb{P}[H \geq a] \leq \mathbb{H}'(a, \mathcal{G}) \text{ for all } H \in \mathcal{G} \text{ with } \mathbb{E}_{\mathbb{P}}[H] \leq 0. \quad (4.9)$$

It follows that

$$\mathbb{P}[H \geq a] \leq \mathbb{H}'(a, \mathcal{G}) \text{ for all } H \in \mathcal{G} \text{ with } \mathbb{E}_{\mathbb{P}}[H] = 0.$$

Since the constraints  $\mathcal{G}$  and the event  $H - \mathbb{E}_{\mathbb{P}}[H] \geq a$  are invariant under scalar translation  $H \mapsto H + c$  it follows that

$$\mathbb{P}[H - \mathbb{E}_{\mathbb{P}}[H] \geq a] \leq \mathbb{H}'(a, \mathcal{G}) \text{ for all } H \in \mathcal{G}.$$

That is, the inequalities (4.8) and (4.9) are equivalent.

McDiarmid's inequality [57, 58] provides the bound  $\mathbb{H}(a, \mathcal{G}) := \exp(-\frac{2a^2}{D^2})$  for (4.8) and its equivalent (4.9), with

$$D^2 := \sum_{i=1}^m D_i^2. \quad (4.10)$$

Define the admissible set corresponding to the assumptions of McDiarmid's inequality:

$$\mathcal{A}_{\text{McD}} = \{(f, \mu) \in \mathcal{G} \times \mathcal{M}^m(\mathcal{X}) \mid \mathbb{E}_{\mu}[f] \leq 0\}, \quad (4.11)$$

and define the optimization problem

$$\mathcal{U}(\mathcal{A}_{\text{McD}}) := \sup_{(f, \mu) \in \mathcal{A}_{\text{McD}}} \mu[f \geq a]. \quad (4.12)$$



Since  $(H, \mathbb{P}) \in \mathcal{A}_{\text{McD}}$  and McDiarmid's inequality  $\mu[f \geq a] \leq \exp(-\frac{2a^2}{D^2})$  is satisfied for all  $(f, \mu) \in \mathcal{A}_{\text{McD}}$ , it follows that

$$\mathbb{P}[H \geq a] \leq \mathcal{U}(\mathcal{A}_{\text{McD}}) \leq \exp\left(-\frac{2a^2}{D^2}\right).$$

Moreover, the inequality on the left is optimal in the sense that, for every  $\varepsilon > 0$ , there exists a McDiarmid-admissible scenario  $(f, \mu)$  satisfying the same assumptions as  $(H, \mathbb{P})$  such that  $\mu[f \geq a] \geq \mathcal{U}(\mathcal{A}_{\text{McD}}) - \varepsilon$ .

To apply the previous results to computing  $\mathcal{U}(\mathcal{A}_{\text{McD}})$ , let  $\mathcal{D} := \{0, 1\}^m$  and define

$$\mathcal{G}_{\mathcal{D}} := \{h \in \mathcal{F}_{\mathcal{D}} \mid \text{Osc}_k(h) \leq D_k \text{ for each } k = 1, \dots, m\},$$

where the inequality  $\text{Osc}_k(h) \leq D_k$  for  $h \in \mathcal{F}_{\mathcal{D}}$  means that

$$|h(s_1, \dots, s_k, \dots, s_m) - h(s_1, \dots, s_{k'}, \dots, s_m)| \leq D_k, \quad (4.13)$$

for all  $s_j \in \{0, 1\}$ ,  $j = 1, \dots, m$ , and all  $s_{k'} \in \{0, 1\}$ . Define the corresponding admissible set

$$\mathcal{A}_{\mathcal{D}} = \{(h, \alpha) \in \mathcal{G}_{\mathcal{D}} \times \mathcal{M}(\{0, 1\}^m) \mid \mathbb{E}_{\alpha}[h] \leq 0\} \quad (4.14)$$

and the optimization problem

$$\mathcal{U}(\mathcal{A}_{\mathcal{D}}) := \sup_{(h, \alpha) \in \mathcal{A}_{\mathcal{D}}} \alpha[h \geq a]. \quad (4.15)$$

**Proposition 4.8.** *It holds that*

$$\mathcal{U}(\mathcal{A}_{\text{McD}}) = \mathcal{U}(\mathcal{A}_{\mathcal{D}}). \quad (4.16)$$

We now provide a further reduction of  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  by reducing  $\mathcal{U}(\mathcal{A}_{\mathcal{D}})$ . To that end, for two vertices  $s$  and  $t$  of  $\mathcal{D} = \{0, 1\}^m$ , let  $I(s, t)$  be the set of indices  $i$  such that  $s_i \neq t_i$ . For  $s \in \mathcal{D}$ , define the function  $h^s \in \mathcal{F}_{\mathcal{D}}$  by

$$h^s(t) = a - \sum_{i \in I(s, t)} D_i.$$

For  $C \subseteq \mathcal{D}$ , define  $h^C \in \mathcal{F}_{\mathcal{D}}$  by

$$h^C(t) := \max_{s \in C} h^s(t) = a - \min_{s \in C} \sum_{i \in I(s, t)} D_i. \quad (4.17)$$

Let  $\mathcal{C} := \{C \mid C \subseteq \mathcal{D}\}$  be the power set of  $\mathcal{D}$  (the set of all subsets of  $\mathcal{D}$ ), define the admissible set  $\mathcal{A}_{\mathcal{C}}$  by

$$\mathcal{A}_{\mathcal{C}} := \{(C, \alpha) \in \mathcal{C} \times \mathcal{M}(\{0, 1\}^m) \mid \mathbb{E}_{\alpha}[h^C] \leq 0\} \quad (4.18)$$

and consider the optimization problem

$$\mathcal{U}(\mathcal{A}_{\mathcal{C}}) := \sup_{(C, \alpha) \in \mathcal{A}_{\mathcal{C}}} \alpha(h^C \geq a). \quad (4.19)$$

**Theorem 4.9.** *It holds that*

$$\mathcal{U}(\mathcal{A}_{\mathcal{D}}) = \mathcal{U}(\mathcal{A}_{\mathcal{C}}). \quad (4.20)$$

**Remark 4.10.** The proof of this reduction theorem utilizes the standard lattice structure of the space of functions  $\mathcal{F}_{\mathcal{D}}$  in a substantial way. To begin with, the reduction to  $\max h = a$  is attained through lattice invariance. Moreover, we have a lattice  $\mathcal{F}_{\mathcal{D}}$ , with sub-lattice  $\mathcal{G}_{\mathcal{D}}$ , and for each  $C \in \mathcal{C}$ , the set  $C_{\mathcal{D}} := \{h \in \mathcal{F}_{\mathcal{D}} \mid \{s \mid h(s) = a\} = C\}$  of functions with value  $a$  precisely on the set  $C$  is a sub-lattice. For a clipped  $h$ , let  $C(h) := \{s \in \mathcal{D} \mid h(s) = a\}$  be the set where  $h$  has the value  $a$ . If for each  $C$  the set

$$\bigcap_{h:C(h)=C} \{f \leq h\} \cap C_{\mathcal{D}} \cap \mathcal{G}_{\mathcal{D}}$$

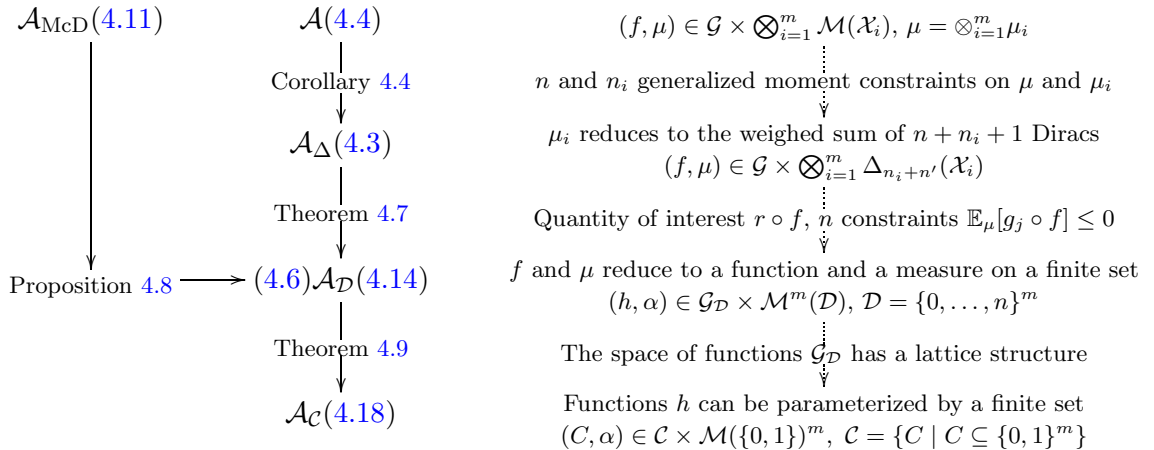
is nonempty, then we obtain a reduction. However, not only is the set nonempty, but the map  $C \mapsto h^C$  is a simple algorithm that produces a point in this intersection, and therefore an explicit reduction. We suspect that the existence of a simple reduction algorithm in this case is due to the lattice structures, and that such structures may be useful in the more general case. Indeed, the condition  $f \leq h$  implies that  $\mathbb{E}_{\alpha}[f] \leq \mathbb{E}_{\alpha}[h]$  for any  $\alpha$ , and the condition that  $\mathbb{E}_{\alpha}[f] \leq \mathbb{E}_{\alpha}[h]$  for all  $\alpha$  implies that  $f \leq h$ , so that the above condition is equivalent to the non-emptiness of

$$\bigcap_{h:C(h)=C} \left\{ \bigcap_{\alpha} \{\mathbb{E}_{\alpha}[f] \leq \mathbb{E}_{\alpha}[h]\} \right\} \cap C_{\mathcal{D}} \cap \mathcal{G}_{\mathcal{D}}.$$

For the more general constraints, we would instead have to solve (i.e. find an element of)

$$\bigcap_{h:C(h)=C} \left\{ \bigcap_{\alpha} \{\mathbf{G}(f, \alpha) \leq \mathbf{G}(h, \alpha)\} \right\} \cap C_{\mathcal{D}} \cap \mathcal{G}_{\mathcal{D}}.$$

**Remark 4.11.** We refer to the following diagram for a summary of the relationships between admissible sets  $\mathcal{A}$ ,  $\mathcal{A}_{\Delta}$ ,  $\mathcal{A}_{\mathcal{D}}$ ,  $\mathcal{A}_{\mathcal{C}}$ ,  $\mathcal{A}_{\text{McD}}$ , the reduction theorems and their assumptions.



## 5 Optimal Concentration Inequalities

In this section, the results of Section 4 will be applied to obtain optimal concentration inequalities under the assumptions of McDiarmid's inequality and Hoeffding's inequality. The following subsection gives explicit concentration results under the assumptions of McDiarmid's inequality, and Subsection 5.2 gives explicit concentration results under the assumptions of Hoeffding's inequality.

Surprisingly, these explicit results show that, although uncertainties may propagate for the true value of  $G$  and  $\mathbb{P}$ , they might not when the information is incomplete on  $G$  and  $\mathbb{P}$ .

We refer to Subsection 10.2 for the proofs of the results in this section.

### 5.1 Explicit solutions under the assumptions of McDiarmid's inequality

In this subsection, we will apply Theorem 4.9 to obtain explicit formulae for the OUQ problem  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  (defined in Equation (4.12)) under the assumptions of McDiarmid's inequality (4.11). More precisely, we will compute  $\mathcal{U}(\mathcal{A}_{\mathcal{C}})$  defined by equation (4.19) and use equalities (4.20) and (4.16) to obtain  $\mathcal{U}(\mathcal{A}_{\text{McD}}) = \mathcal{U}(\mathcal{A}_{\mathcal{C}})$ . Observe that all the following optimization problems possess solutions because they involve the optimization of a continuous function (with respect to  $\alpha$ ) in a compact space.

Since the inequalities (4.8) and (4.9) are equivalent, it follows that

$$\mathcal{U}(\mathcal{A}_{\text{McD}}) = \sup_{(f, \mu) \in \mathcal{G} \times \mathcal{M}_m} \mu[f \geq a + \mathbb{E}_{\mu}[f]].$$

In particular, if  $\mathbb{E}_{\mu}[f] \leq 0$  is replaced by  $\mathbb{E}_{\mu}[f] \leq b$  or  $\mathbb{E}_{\mu}[f] = b$  in McDiarmid's inequality assumptions (4.11), then the results given in this section remain valid by replacing  $a$  by  $M := a - b$  (observe that  $M$  plays the role of a margin).

Those results should be compared with McDiarmid's inequality [57, 58], which provides the bound

$$\sup_{(f, \mu) \in \mathcal{G} \times \mathcal{M}_m} \mu[f \geq a + \mathbb{E}_{\mu}[f]] \leq \exp\left(-\frac{2a^2}{\sum_{i=1}^m D_i^2}\right). \quad (5.1)$$

The statements of the theorem will be given assuming that  $a \geq 0$ ; in the complementary case of  $a < 0$ , the solution is simply  $\mathcal{U}(\mathcal{A}_{\text{McD}}) = 1$ .

To the best of the authors' knowledge, the optimal bounds given here are new. There is a substantial literature relating to optimization of concentration bounds and de-randomization algorithms (see for instance [85] and references therein) but, as far as the authors know, those bounds were suboptimal because they were obtained through the moment generating function technique.

### 5.1.1 Explicit solutions in dimensions one and two

**Theorem 5.1** (Explicit solution for  $m = 1$ ). *For  $m = 1$ ,  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  is given by*

$$\mathcal{U}(\mathcal{A}_{\text{McD}}) = \begin{cases} 0, & \text{if } D_1 \leq a, \\ 1 - \frac{a}{D_1}, & \text{if } 0 \leq a \leq D_1. \end{cases} \quad (5.2)$$

**Theorem 5.2** (Explicit solution for  $m = 2$ ). *For  $m = 2$ ,  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  is given by*

$$\mathcal{U}(\mathcal{A}_{\text{McD}}) = \begin{cases} 0, & \text{if } D_1 + D_2 \leq a, \\ \frac{(D_1 + D_2 - a)^2}{4D_1D_2}, & \text{if } |D_1 - D_2| \leq a \leq D_1 + D_2, \\ 1 - \frac{a}{\max(D_1, D_2)}, & \text{if } 0 \leq a \leq |D_1 - D_2|. \end{cases} \quad (5.3)$$

See Sub-figures 5.2(a), 5.2(b) and 5.2(c) for illustrations comparing the McDiarmid and OUQ bounds for  $m = 2$  (as functions of  $(D_1, D_2)$ , with mean performance 0 and failure threshold  $a = 1$ , the OUQ bound is calculated using the explicit solution (5.3)). Observe that

- If  $a \leq D_1 - D_2$ , then a decrease in  $D_2$  does not lead to a decrease in the OUQ bound  $\mathcal{U}(\mathcal{A}_{\text{McD}})$ . In other words, if most of the uncertainty is contained in the first variable ( $a + D_2 \leq D_1$ ), then the uncertainty associated with the second variable does not affect the global uncertainty; a reduction of the global uncertainty requires a reduction in  $D_1$ .
- For  $D_1 + D_2 = 2a$ , the ratio between the OUQ bound and the McDiarmid bound is minimized near the diagonal.

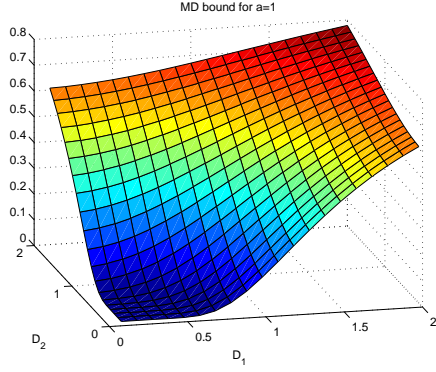
**Remark 5.3.** The maximum of (5.3) over  $D_1, D_2$  under the constraints  $D_1 + D_2 = D$  and  $D_1 \geq D_2$  is achieved at  $D_2 = 0$  and is equal to  $1 - a/D$ . The minimum of (5.3) over  $D_1, D_2$  under the constraints  $D_1 + D_2 = D$  and  $D_1 \geq D_2$  is achieved on the diagonal  $D_1 = D_2$  and is equal to  $(1 - a/D)^2$ .

### 5.1.2 Explicit solution in dimension three

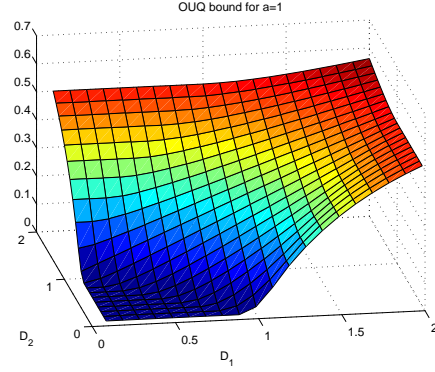
Assume that  $D_1 \geq D_2 \geq D_3$ . Write

$$\mathcal{F}_1 := \begin{cases} 0, & \text{if } D_1 + D_2 + D_3 \leq a, \\ \frac{(D_1 + D_2 + D_3 - a)^3}{27D_1D_2D_3}, & \text{if } D_1 + D_2 - 2D_3 \leq a \leq D_1 + D_2 + D_3, \\ \frac{(D_1 + D_2 - a)^2}{4D_1D_2}, & \text{if } D_1 - D_2 \leq a \leq D_1 + D_2 - 2D_3, \\ 1 - \frac{a}{D_1}, & \text{if } 0 \leq a \leq D_1 - D_2. \end{cases} \quad (5.4)$$

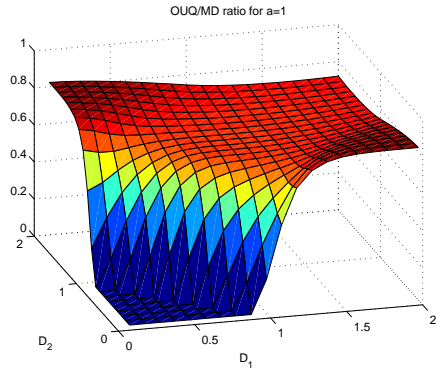
Figure 5.1: Comparison of the McDiarmid and OUQ bounds with zero mean performance and failure threshold  $a = 1$ .



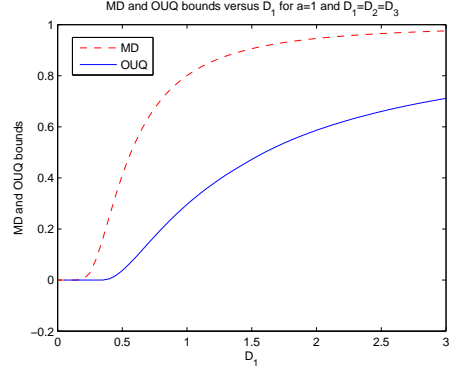
(a) McDiarmid upper bound,  $m = 2$



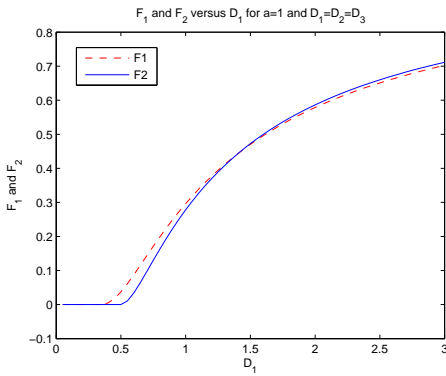
(b) OUQ upper bound,  $m = 2$



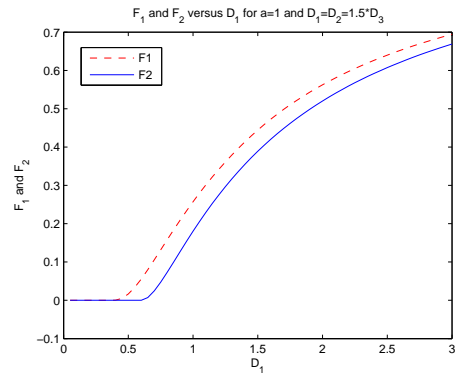
(c) Ratio of the two bounds: OUQ bound divided by McDiarmid bound,  $m = 2$



(d) McDiarmid vs OUQ bound,  $m = 3$  and  $D_1 = D_2 = D_3$



(e)  $\mathcal{F}_1$  vs  $\mathcal{F}_2$ ,  $D_1 = D_2 = D_3$



(f)  $\mathcal{F}_1$  vs  $\mathcal{F}_2$ ,  $D_1 = D_2 = \frac{3}{2}D_3$

and

$$\mathcal{F}_2 := \max_{i \in \{1,2,3\}} \phi(\gamma_i) \psi(\gamma_i) \quad (5.5)$$

where

$$\psi(\gamma) := \gamma^2 \left( 2 \frac{D_2}{D_3} - 1 \right) - 2\gamma \left( 3 \frac{D_2}{D_3} - 1 \right) + \frac{\gamma}{1+\gamma} \left( 8 \frac{D_2}{D_3} - 2 \frac{a}{D_3} \right)$$

and  $\gamma_1, \gamma_2, \gamma_3$  are the roots (in  $\gamma$ ) of the cubic polynomial

$$(1+\gamma)^3 - A(1+\gamma)^2 + B = 0, \quad (5.6)$$

where

$$A := \frac{5D_2 - 2D_3}{2D_2 - D_3} \quad \text{and} \quad B := \frac{4D_2 - a}{2D_2 - D_3}.$$

Define a function  $\phi$  by

$$\phi(\gamma) := \begin{cases} 1, & \text{if } \gamma \in (0, 1) \text{ and } \theta(\gamma) \in (0, 1), \\ 0, & \text{otherwise,} \end{cases}$$

where

$$\theta(\gamma) := 1 - \frac{a}{D_3(1-\gamma^2)} + \frac{D_2}{D_3} \frac{1-\gamma}{1+\gamma}.$$

By the standard formula for the roots of a cubic polynomial, the roots of (5.6) are given by

$$\begin{aligned} \gamma_1 &:= -1 - \frac{1}{3}(-A + \kappa_1 + \kappa_2), \\ \gamma_2 &:= -1 - \frac{1}{3}(-A + \omega_2 \kappa_1 + \omega_1 \kappa_2), \\ \gamma_3 &:= -1 - \frac{1}{3}(-A + \omega_1 \kappa_1 + \omega_2 \kappa_2), \end{aligned}$$

where

$$\begin{aligned} \omega_1 &:= -\frac{1}{2} + \frac{\sqrt{3}}{2}i, \quad \omega_2 := -\frac{1}{2} - \frac{\sqrt{3}}{2}i, \quad \kappa_1 := \left( \frac{\beta_1 + \sqrt{\beta_2}}{2} \right)^{\frac{1}{3}}, \\ \kappa_2 &:= \left( \frac{\beta_1 - \sqrt{\beta_2}}{2} \right)^{\frac{1}{3}}, \quad \beta_1 := -2A^3 + 27B \quad \text{and} \quad \beta_2 := \beta_1^2 - 4A^6. \end{aligned}$$

Since there are 3 possible values for each cube root,  $\kappa_1$  and  $\kappa_2$  must be taken so that they satisfy  $\kappa_1 \kappa_2 = A^2$ .

**Theorem 5.4** (Explicit solution for  $m = 3$ ). *For  $m = 3$  with  $D_1 \geq D_2 \geq D_3$ ,  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  is given by*

$$\mathcal{U}(\mathcal{A}_{\text{McD}}) = \max(\mathcal{F}_1, \mathcal{F}_2). \quad (5.7)$$

**Remark 5.5.** Sub-figure 5.2(d) compares the McDiarmid and OUQ bounds for  $m = 3$ , with zero mean performance,  $D_1 = D_2 = D_3$ , and failure threshold  $a = 1$ . Sub-figure 5.2(e) shows that that  $\mathcal{F}_2 > \mathcal{F}_1$  for  $D_1$  large enough. Sub-figure 5.2(f) shows that if  $D_1 = D_2 = \frac{3}{2}D_3$ , then  $\mathcal{F}_2 < \mathcal{F}_1$  for all  $D_1$ . Therefore, Sub-figures 5.2(e) and 5.2(f) suggest that the inequality  $\mathcal{F}_2 > \mathcal{F}_1$  holds only if  $D_3 \approx D_2$ , and  $D_2$  is large enough relative to  $D_1$ .

**Remark 5.6.** For the application to the (SPHIR facility) admissible set (1.9) (described in Subsection 1.2.2), the sub-diameters of the surrogate  $H$  are:  $8.86 \text{ mm}^2$  for thickness ( $D_1$ ),  $7.20 \text{ mm}^2$  for velocity ( $D_2$ ), and  $4.17 \text{ mm}^2$  for obliquity ( $D_3$ ). These values have been obtained by solving the optimization problems defined by (1.10) with  $f = H$  and  $i = 1, 2, 3$ . The application of Theorem 5.4 with these sub-diameters and  $a = 5.5 \text{ mm}^2$  leads to  $\mathcal{F}_2 = 0.253$  and  $\mathcal{F}_1 = 0.437$  (see (5.4) and (5.5) for the definition and interpretation of  $\mathcal{F}_1$  and  $\mathcal{F}_2$ ). In particular, since  $D_1 - D_2 \leq a \leq D_1 + D_2 - 2D_3$ , it follows from (5.4) that the obliquity sub-diameter does not impact  $\mathcal{F}_1$  (decreasing  $D_3$  down to zero does not change the optimal bound 43.7% obtained from the third line of (5.4)).

### 5.1.3 Solution in dimension $m$

For  $C_0 \in \mathcal{C}$ , write

$$\mathcal{U}(\mathcal{A}_{C_0}) = \sup_{\alpha: (C_0, \alpha) \in \mathcal{A}_{\mathcal{C}}} \alpha[h^{C_0} \geq a], \quad (5.8)$$

where  $h^{C_0}$  is defined by equation (4.17).

**Proposition 5.7.** Assume that  $D_1 \geq \dots \geq D_{m-1} \geq D_m$ . For  $C_0 := \{(1, 1, \dots, 1, 1)\}$ , it holds that

$$\mathcal{U}(\mathcal{A}_{C_0}) = \begin{cases} 0, & \text{if } \sum_{j=1}^m D_j \leq a, \\ \frac{(\sum_{j=1}^m D_j - a)^m}{m^m \prod_{j=1}^m D_j}, & \text{if } \sum_{j=1}^m D_j - mD_m \leq a \leq \sum_{j=1}^m D_j, \\ \frac{(\sum_{j=1}^k D_j - a)^k}{k^k \prod_{j=1}^k D_j}, & \text{if, for } k \in \{1, \dots, m-1\}, \\ & \sum_{j=1}^k D_j - kD_k \leq a \leq \sum_{j=1}^{k+1} D_j - (k+1)D_{k+1}. \end{cases} \quad (5.9)$$

**Remark 5.8.** The maximum of (5.9) over  $D_1, \dots, D_m$  under the constraints  $D_1 + \dots + D_m = D$  and  $D_1 \geq \dots \geq D_m$  is achieved at  $D_1 = D$  and is equal to  $1 - a/D$ . The minimum of (5.9) over  $D_1, \dots, D_m$  under the constraints  $D_1 + \dots + D_m = D$  and  $D_1 \geq \dots \geq D_m$  is achieved on the diagonal  $D_1 = \dots = D_m$  and is equal to  $(1 - a/D)^m$ .

**Proposition 5.9.** Assume that  $D_1 \geq \dots \geq D_{m-1} \geq D_m$ . If  $a \geq \sum_{j=1}^{m-2} D_j + D_m$ , then  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  is given by equation (5.9).

**Remark 5.10.** It follows from the previous proposition that, in arbitrary dimension  $m$ , the tail of  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  with respect to  $a$  is given by (5.9). Although we do not have an analytic solution for  $m \geq 4$  and  $a < \sum_{j=1}^{m-2} D_j + D_m$ , a numerical solution can be

obtained by solving the finite-dimensional optimization problem (4.19) with variables  $(C, \alpha)$ . Observe that the range of  $\alpha$  is  $[0, 1]^m$ . Although the range of  $C$  is the set of subsets of  $\{0, 1\}^m$ , we conjecture (based on symmetry and monotonicity arguments) that the extremum of (4.19) can be achieved by restricting  $C$  to sets  $C_q$  defined by  $\{s \in [0, 1]^m \mid \sum_{i=1}^m s_i \geq q \text{ (with } q \in \{1, \dots, m\})\}$ .

## 5.2 Explicit solutions under the assumptions of Hoeffding's inequality

This subsection treats a further special case of OUQ, where the assumptions are those of Hoeffding's inequality [35]. Define the admissible set

$$\mathcal{A}_{\text{Hfd}} := \left\{ (f, \mu) \left| \begin{array}{l} f = X_1 + \dots + X_m, \\ \mu \in \bigotimes_{i=1}^m \mathcal{M}([b_i - D_i, b_i]), \\ \mathbb{E}_\mu[f] \leq 0 \end{array} \right. \right\}, \quad (5.10)$$

and define the optimization problem

$$\mathcal{U}(\mathcal{A}_{\text{Hfd}}) := \sup_{(f, \mu) \in \mathcal{A}_{\text{Hfd}}} \mu[f \geq a].$$

By Hoeffding's inequality, for  $a \geq 0$ ,

$$\mathcal{U}(\mathcal{A}_{\text{Hfd}}) \leq \exp \left( -2 \frac{a^2}{\sum_i^m D_i^2} \right).$$

**Theorem 5.11.** *If  $m = 2$ , then*

$$\mathcal{U}(\mathcal{A}_{\text{Hfd}}) = \mathcal{U}(\mathcal{A}_{\text{McD}}). \quad (5.11)$$

**Remark 5.12.** Another proof of Theorem 5.11 can be obtained using entirely different methods than presented in Section 10.2. Although omitted for brevity, this method may be useful in higher dimensions, so we describe an outline of it here. We begin at the reduction obtained through Proposition 4.8 to the hypercube. Whereas the proof of Theorem 5.11 first applies the reduction of Theorem 4.9 to subsets of the hypercube, here we instead fix the oscillations in each direction to be  $0 \leq d_i \leq D_i$ , and solve the fixed  $d := (d_1, d_2)$  case, not using a Lagrangian-type analysis but a type of spectral reduction. We then show that the resulting value  $\mathcal{U}(d)$  is increasing in  $d$  with respect to the standard (lexicographic) partial order on vectors. The result then easily follows by taking the supremum over all vectors  $0 \leq d \leq D$ .

**Theorem 5.13.** *Let  $m = 3$ , and define  $\mathcal{F}_1$  and  $\mathcal{F}_2$  as in Theorem 5.4. If  $\mathcal{F}_1 \geq \mathcal{F}_2$ , then*

$$\mathcal{U}(\mathcal{A}_{\text{Hfd}}) = \mathcal{U}(\mathcal{A}_{\text{McD}}). \quad (5.12)$$

*If  $\mathcal{F}_1 < \mathcal{F}_2$ , then*

$$\mathcal{U}(\mathcal{A}_{\text{Hfd}}) < \mathcal{U}(\mathcal{A}_{\text{McD}}). \quad (5.13)$$



Under the assumptions of Hoeffding’s inequality, each variable  $X_i$  is bounded from below and from above. Without the upper bounds on the variables  $X_i$ , it is possible to use additional reduction properties and conjecture an explicit form for the optimal inequality on  $\mu[X_1 + \dots + X_m \geq a]$ . Here we refer to the work and conjecture of Samuels [77] (see also [42, p.542]), which has been proven true for  $m = 1, 2, 3$ .

**Remark 5.14.** The optimal Hoeffding inequality can be used for additive models (with response functions of the form  $X_1 + \dots + X_m$ ) but also to obtain optimal probabilities of deviations for empirical means. Furthermore the fact that the optimal concentration inequalities corresponding to Hoeffding’s or McDiarmid’s assumptions are the same for  $m = 2$  and possibly distinct for  $m = 3$  is a simple but fundamental result analogous to Stein’s paradox [25].

## 6 Computational Implementation

In this section, we discuss the numerical implementation of OUQ algorithms for the analytical surrogate model for hypervelocity impact introduced in Subsection 1.2.

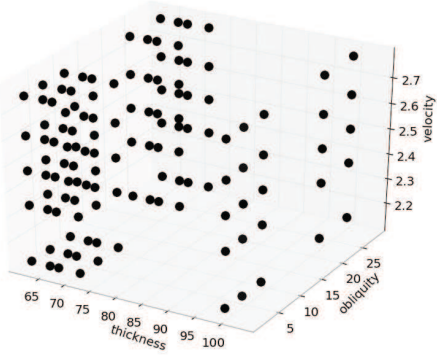
### 6.1 Extreme points of reduced OUQ problems are attractors

We consider again the computation of the optimal bound  $\mathcal{U}(\mathcal{A}_H)$  (where  $\mathcal{A}_H$  is the information set given by Equation (1.8)) via the identity (1.12) derived from the reduction results of Section 4. For  $\#\text{supp}(\mu_i) \leq 2$ ,  $i = 1, 2, 3$ , Figure 1.3 has shown that numerical simulations collapse to two-point support. Figure 6.1 shows that, even when a wider search is performed (i.e., over measures  $\mu \in \bigotimes_{i=1}^3 \Delta_k(\mathcal{X}_i)$  for  $k > 1$ ), the calculated maximizers for these problems maintain two-point support: the velocity and obliquity marginals each collapse to a single Dirac mass, and the plate thickness marginal collapses to have support on the two extremes of its range. As expected, optimization over a larger search space is more computationally intensive and takes longer to perform. This observation suggests that the extreme points of the reduced OUQ problems are, in some sense, attractors — this point will be revisited in the next subsection.

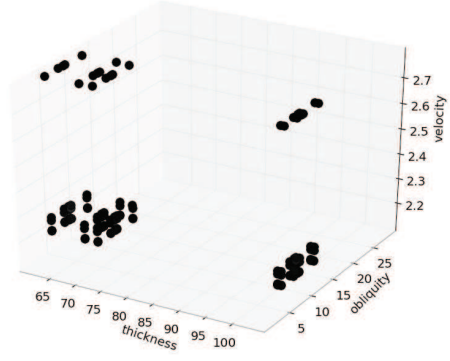
We also refer to Figures 1.4 and 6.2 for plots of the locations and weights of the Dirac masses forming each marginal  $\mu_i$  as functions of the number of iterations. Note that the lines for *thickness* and *thickness weight* are of the same color if they correspond to the same support point for the measure. In particular, Figure 6.2 shows that at iteration number 3500 the *thickness* support point at 62.5 mils (shown in Figure 6.1) has zero weight.

### 6.2 Coagulation–Fragmentation algorithm for OUQ

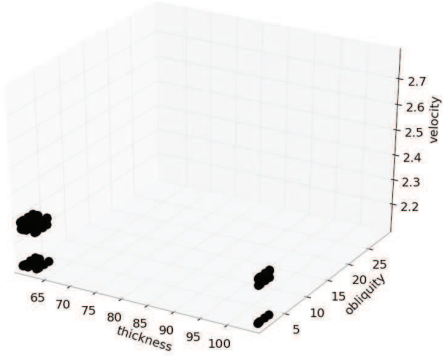
The results of Sections 4 and 5 give explicit a priori bounds on the number of Dirac masses sufficient to find the lower and upper bounds  $\mathcal{L}(\mathcal{A})$  and  $\mathcal{U}(\mathcal{A})$  when the admissible set  $\mathcal{A}$  is given by independence and linear inequality constraints. However, it is possible that reduction properties are present for more general admissible sets  $\mathcal{A}$ . Can such



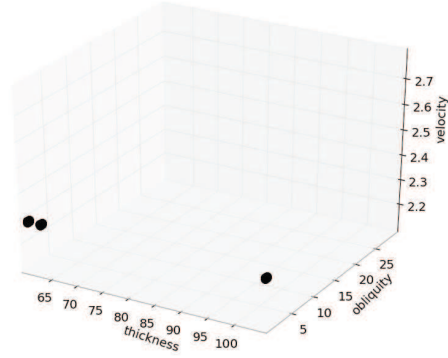
(g) support points at iteration 0



(h) support points at iteration 1000



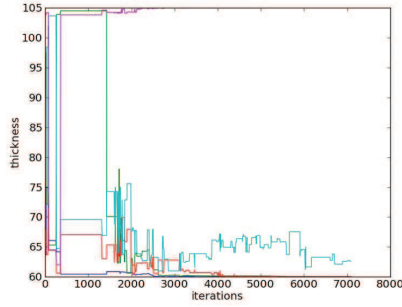
(i) support points at iteration 3000



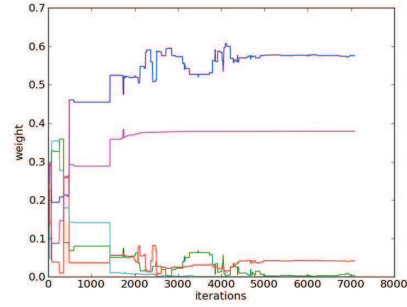
(j) support points at iteration 7100

Figure 6.1: For  $\#\text{supp}(\mu_i) \leq 5$ ,  $i = 1, 2, 3$ , the maximizers of the OUQ problem (1.12) associated with the information set (1.8) collapse to two-point support. Velocity, obliquity and plate thickness marginals collapse as in Figure 1.3. At iteration 7100, the thickness support point at 62.5 mils has zero weight, as can be seen in Figure 6.2.

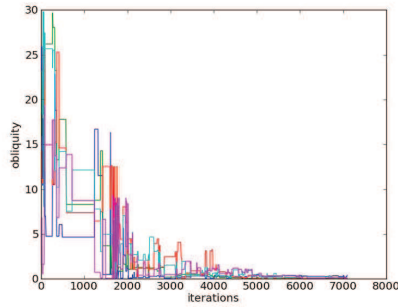
Figure 6.2: Time evolution of the genetic algorithm search for the OUQ problem (1.12) associated with the information set (1.8) for  $\#\text{supp}(\mu_i) \leq 5$  for  $i = 1, 2, 3$ , as optimized by *mystic*. Four of the five thickness support points quickly converge to the extremes of its range, with weights 0.024, 0.058, and 0.539 at 60 mils and weight 0.379 at 105 mils. The thickness support point that does not converge to an extremum has zero weight. Obliquity and velocity each collapse to a single support point, again with the corresponding weights demonstrating fluctuations due to degeneracies.



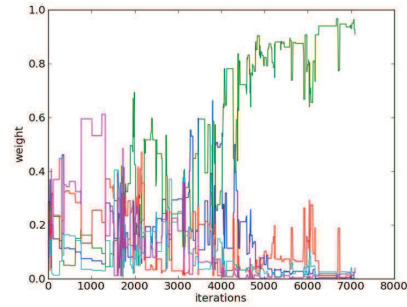
(a) convergence for thickness



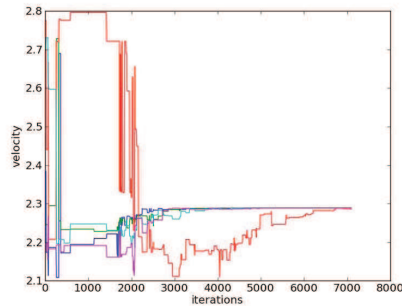
(b) convergence for thickness weight



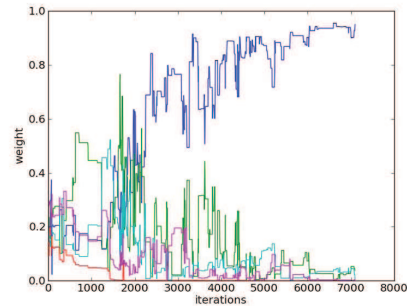
(c) convergence for obliquity



(d) convergence for obliquity weight



(e) convergence for velocity



(f) convergence for velocity weight

“hidden” reduction properties be detected by computational means, even in the absence of theorems that prove their existence?

Consider again the results of the previous subsection. Theorem 4.1 provides an a priori guarantee that, to find  $\mathcal{U}(\mathcal{A})$ , it is sufficient to search the reduced feasible set  $\mathcal{A}_\Delta$ , which consists of those  $\mu \in \mathcal{A}$  whose marginal distributions each have support on at most two points. However, Figure 6.1 provides numerical evidence that something much stronger is true: even if we search among measures  $\mu \in \bigotimes_{i=1}^3 \Delta_k(\mathcal{X}_i)$  for  $k \geq 1$ , the measures collapse to an optimizer

$$\mu^* \in \Delta_1(\mathcal{X}_1) \otimes \Delta_0(\mathcal{X}_2) \otimes \Delta_0(\mathcal{X}_3)$$

(that is, two-point support on the thickness axis, and one-point support on the obliquity and velocity axes). In some sense, the (small support) optimizers are attractors for the optimization process even when the optimization routine is allowed to search over measures with larger support than that asserted by Theorem 4.1.

Therefore, we propose the following general algorithm for the detection of *hidden* reduction properties. Let an admissible set  $\mathcal{A}$  be given; for  $k \in \mathbb{N}$ , let

$$\mathcal{A}_k := \{(f, \mu) \in \mathcal{A} \mid \mu \in \Delta_k(\mathcal{X})\}$$

be the collection of admissible scenarios such that  $\mu$  has support on at most  $k+1$  points of  $\mathcal{X}$ .

1. Fix any initial value of  $k \in \mathbb{N}$ .
2. Numerically calculate  $\mathcal{U}(\mathcal{A}_k)$  and obtain a numerical (approximate) maximizer  $\mu^* \in \mathcal{A}_k$ .
3. Calculate  $\#\text{supp}(\mu^*)$  and proceed as follows:
  - If  $\#\text{supp}(\mu^*) < k+1$ , then the measure has *coagulated* to have less-than-maximally-sized support and we terminate the algorithm.
  - If  $\#\text{supp}(\mu^*) = k+1$ , then no coagulation/reduction has yet been observed. We enter a *fragmentation* phase: replace  $k$  by any  $k' > k$  and return to step 2.

**Remark 6.1.** It would be more accurate to say that the above algorithm is a sketch of an algorithm, and that its details should be adjusted to fit the circumstances of application. For example, if the admissible set  $\mathcal{A}$  includes an independence constraint, then it would be appropriate to base decisions upon the cardinality of the support of the *marginal* distributions of  $\mu^*$ , not on the cardinality of the support of  $\mu^*$  itself. The termination of the algorithm if  $\#\text{supp}(\mu^*) < k+1$  is motivated by supposition that a *hidden* reduction property has been found and that  $\mathcal{U}(\mathcal{A})$  has an (approximate) optimizer in  $\mathcal{A}_k$ .

**Remark 6.2.** We reiterate the point made in Remark 4.3 that these methods apply to more general situations than finite convex combinations of Dirac measures; finite convex combinations of Dirac measures are simply a well-known class of geometrically extreme probability measures (with respect to which numerical integration happens to be very easy), and can be replaced by the extremal points of any class of probability measures

as required by the situation of study. For example, if the OUQ problem of interest involved the invariant measures for some measurable transformation  $T: \mathcal{X} \rightarrow \mathcal{X}$ , then each occurrence of  $\Delta_k(\mathcal{X})$  above would be replaced by

$$\mathcal{E}_k^T(\mathcal{X}) := \left\{ \sum_{j=0}^k \alpha_j \mu_j \left| \begin{array}{l} \text{for each } j = 0, \dots, k, \alpha_j \geq 0, \\ \mu_j \in \mathcal{M}(\mathcal{X}) \text{ is ergodic with respect to } T, \\ \text{and } \sum_{j=0}^k \alpha_j = 1 \end{array} \right. \right\}.$$

### 6.3 The OUQ algorithm in the *mystic* framework

As posed above, OUQ at the high level is a global optimization of a cost function that satisfies a set of constraints. This optimization is performed in *mystic* using the differential evolution algorithm of Price & Storn [72, 88], with constraints satisfied through a modified Lagrange multiplier method [60].

The *mystic* optimization framework [59] provides a collection of optimization algorithms and tools that lowers the barrier to solving complex optimization problems. Specifically, *mystic* provides flexibility in specifying the optimization algorithm, constraints, and termination conditions. For example, *mystic* classifies constraints as either “bounds constraints” (linear inequality constraints that involve precisely one input variable) or “non-bounds constraints” (constraints between two or more parameters), where either class of constraint modifies the cost function accordingly in attempt to maximize algorithm accuracy and efficiency. Every *mystic* optimizer provides the ability to apply bounds constraints generically and directly to the cost function, so that the difference in the speed of bounds-constrained optimization and unconstrained optimization is minimized. *Mystic* also enables the user to impose an arbitrary parameter constraint function on the input of the cost function, allowing non-bounds constraints to be generically applied in any optimization problem.

The *mystic* framework was extended for the OUQ algorithm. A modified Lagrange multiplier method was added to *mystic*, where an internal optimization is used to satisfy the constraints at each iteration over the cost function [60]. Since evaluation of the cost function is commonly the most expensive part of the optimization, our implementation of OUQ in *mystic* attempts to minimize the number of cost function evaluations required to find an acceptable solution. By satisfying the constraints within some tolerance at each iteration, our OUQ algorithm will (likely) numerically converge much more quickly than if we were to apply constraints by invalidating generated results (i.e. filtering) at each iteration. In this way, we can use *mystic* to efficiently solve for rare events, because the set of input variables produced by the optimizer at each iteration will also be an admissible point in problem space — this feature is critical in solving OUQ problems, as tens of thousands of function evaluations may be required to produce a solution. We refer to [60] for a detailed description of the implementation of the OUQ algorithm in the *mystic* framework (we also refer to [61]).

**Remark 6.3.** Our implementation of the OUQ algorithm in *mystic* utilizes a nested optimization (an inner loop) to solve an arbitrary set of parameter constraints at each

evaluation of the cost function. We use evolutionary algorithms because they are robust and especially suited to the inner loop (i.e., at making sure that the constraints are satisfied, local methods and even some global method are usually not good enough for this). We also note that the outer loop can be relaxed to other methods (leading to a reduction in the total number of function evaluations by an order of magnitude). Finally, although we observe approximate extremizers that are “computationally” distinct (Figure 6.2 shows that mass is traded wildly between practically coincident points), we have not observed yet “mathematically” distinct extrema.

**Measures as data objects.** Theorem 4.1 states that a solution to an OUQ problem, with linear constraints on marginal distributions, can be expressed in terms of products of convex linear combinations of Dirac masses. In our OUQ algorithm, the optimizer’s parameter generator produces new parameters each iteration, and hence produces new product measures to be evaluated within the cost function. For instance, the response function  $H$ , as defined by  $H(h, \theta, v)$  in (1.5), requires a product measure of dimension  $n = 3$  for support. In Example (1.8), the mean perforation area is limited to  $[m_1, m_2] = [5.5, 7.5] \text{ mm}^2$ , the parameters  $h, \theta, v$  are bounded by the range provided by (1.7), and products of convex combinations of Dirac masses are used as the basis for support. The corresponding OUQ code parameterizes the Dirac masses by their weights and positions.

More generally, it is worth noting that our computational implementation of OUQ is expressed in terms of methods that act on a hierarchy of parameterized measure data objects. Information is thus passed between the different elements of the OUQ algorithm code as a list of parameters (as required by the optimizer) or as a parameterized measure object. *Mystic* includes methods to automate the conversion of measure objects to parameter lists and vice versa, hence the bulk of the OUQ algorithm code (i.e. an optimization on a product measure) is independent of the selection of basis of the product measure itself. In particular, since the measure data objects can be decoupled from the rest of the algorithm, the product measure representation can be chosen to best provide support for the model, whether it be a convex combination of Dirac masses as required by Example (1.8), or measures composed of another basis such as Gaussians. More precisely, this framework can naturally be extended to Gaussians merely by adding covariance matrices as data object variables and by estimating integral moments equations (with a Monte Carlo method for instance).

## 7 Application to the Seismic Safety Assessment of Structures

In this section, we assess the feasibility of the OUQ formalism by means of an application to the safety assessment of truss structures subjected to ground motion excitation. This application contains many of the features that both motivate and challenge UQ, including imperfect knowledge of random inputs of high dimensionality, a time-dependent and complex response of the system, and the need to make high-consequence decisions pertaining to the safety of the system. The main objective of the analysis is to assess the

safety of a structure knowing the maximum magnitude and focal distance of the earthquakes that it may be subjected to, with limited information and as few assumptions as possible.

## 7.1 Formulation in the time domain

### 7.1.1 Formulation of the problem

For definiteness, we specifically consider truss structures undergoing a purely elastic response, whereupon the vibrations of the structure are governed by the structural dynamics equation

$$M\ddot{u}(t) + C\dot{u}(t) + Ku(t) = f(t), \quad (7.1)$$

where  $u(t) \in \mathbb{R}^N$  collects the displacements of the joints,  $M$  is the mass matrix,  $C$  is the damping matrix,  $K$  is the stiffness matrix and  $f(t) \in \mathbb{R}^N$  are externally applied forces, such as dead-weight loads, wind loads and others. The matrices  $M$ ,  $C$  and  $K$  are of dimension  $N \times N$ , symmetric and strictly positive definite. Let  $T$  be an  $N \times 3$  matrix such that:  $T_{ij} = 1$  if the  $i$ th degree-of-freedom is a displacement in the  $j$ th coordinate direction; and  $T_{ij} = 0$  otherwise. In addition, let  $u_0(t) \in \mathbb{R}^3$  be a ground motion. Then,  $Tu_0(t)$  represents the motion obtained by translating the entire structures rigidly according to the ground motion. We now introduce the representation

$$u(t) = Tu_0(t) + v(t), \quad (7.2)$$

where  $v(t)$  now describes the vibrations of the structure relative to its translated position. Inserting (7.2) into (7.1) and using  $KT = 0$  and  $CT = 0$  (implied by translation invariance), we obtain

$$M\ddot{v}(t) + C\dot{v}(t) + Kv(t) = f(t) - MT\ddot{u}_0(t), \quad (7.3)$$

where  $-MT\ddot{u}_0(t)$  may be regarded as the effective forces induced in the structure by the ground motion (we start from rest). We shall assume that the structure is required to remain in the elastic domain for purposes of certification. Suppose that the structure has  $J$  members and that all the external loads are applied to the joints of the structure. Let  $L$  be a  $J \times N$  matrix such that the entries of the vector  $Lv$  give the axial strains of the members. The certification condition is, therefore,

$$\|L_i v\|_\infty < S_i, \quad i = 1, \dots, J, \quad (7.4)$$

where  $S_i$  is the yield strain of the  $i$ th member and  $\|f\|_\infty := \text{ess sup } |f|$  is the  $L^\infty$ -norm of a function  $f: \mathbb{R} \rightarrow \mathbb{R}$ . In what follows we will write

$$Y_i = L_i v \quad i = 1, \dots, J, \quad (7.5)$$

for the member strains. Due to the linearity of the structure, a general solution of (7.3) may be formally obtained by means of a modal analysis. Thus, let  $q_\alpha \in \mathbb{R}^N$  and  $\omega_\alpha > 0$ ,

$\alpha = 1, \dots, N$ , be the eigenvectors and eigenfrequencies corresponding to the symmetric eigenvalue problem  $(K - \omega_\alpha^2 M)q_\alpha = 0$ , normalized by  $q_\alpha^T M q_\alpha = 1$ . Let

$$v(t) = \sum_{\alpha=1}^N v_\alpha(t) q_\alpha \quad (7.6)$$

be the modal decomposition of  $v(t)$ . Using this representation, the equation of motion (7.3) decomposes into the modal equations

$$\ddot{v}_\alpha(t) + 2\zeta_\alpha \omega_\alpha \dot{v}_\alpha(t) + \omega_\alpha^2 v_\alpha(t) = q_\alpha^T (f(t) - MT\ddot{u}_0(t)), \quad (7.7)$$

where we have assumed that the eigenmodes  $q_\alpha$  are also eigenvectors of  $C$  and  $\zeta_\alpha$  is the damping ratio for mode  $i$ . The solution of (7.7) is given by the hereditary integral

$$v_\alpha(t) = - \int_0^t e^{-\zeta_\alpha \omega_\alpha (t-\tau)} \sin[\omega_\alpha (t-\tau)] (q_\alpha^T MT\ddot{u}_0(\tau)) \frac{d\tau}{\omega_\alpha}, \quad (7.8)$$

where, for simplicity, we set  $f = 0$  and assume that the structure starts from rest and without deformation at time  $t = 0$ . We can now regard structures oscillating under the action of a ground motion as systems that take the ground motion acceleration  $\ddot{u}_0(t)$  as input and whose performance measures of interest are the member strains  $Y_i$ ,  $i = 1, \dots, J$ . The response function  $F$  mapping the former to the latter is given by composing (7.8), (7.6) and (7.5).

### 7.1.2 Formulation of the information set

In order to properly define the certification problem we proceed to define constraints on the inputs, i.e. the information set associated with the ground motion acceleration. As in [87], we regard the ground motion at the site of the structure as a combination of two factors: the earthquake source  $s$  and the earth structure through which the seismic waves propagate; this structure is characterized by a transfer function  $\psi$ . Let  $\star$  denote the convolution operator; the ground motion acceleration is then given by

$$\ddot{u}_0(t) := (\psi \star s)(t). \quad (7.9)$$

We assume that  $s$  is a sum of boxcar time impulses (see [87] page 230) whose amplitudes and durations are random, independent, not identically distributed and of unknown distribution. More precisely, we assume that

$$s(t) := \sum_{i=1}^B X_i s_i(t), \quad (7.10)$$

where  $X_1, \dots, X_B$  are independent (not necessarily identically distributed) random variables with unknown distribution with support in  $[-a_{\max}, a_{\max}]^3$  ( $s_i$ ,  $B$  and  $a_{\max}$  are defined below) and such that  $\mathbb{E}[X_i] = 0$ . We also assume the components  $(X_{i,1}, X_{i,2}, X_{i,3})$



of the vectors  $X_i$  to be independent. Since we wish to bound the probability that a structure will fail when it is struck by an earthquake of magnitude  $M_L$  in the Richter (local magnitude) scale and hypocentral distance  $R$ , we adopt the semi-empirical expression proposed by Esteva [28] (see also [63]) for the maximum ground acceleration

$$a_{\max} := \frac{a_0 e^{\lambda M_L}}{(R_0 + R)^2}, \quad (7.11)$$

where  $a_0$ ,  $\lambda$  and  $R_0$  are constants. For earthquakes on firm ground, Esteva [28] gives  $a_0 = 12.3 \cdot 10^6 \text{ m}^3 \cdot \text{s}^{-2}$ ,  $\lambda = 0.8$  and  $R_0 = 25 \cdot 10^3 \text{ m}$ .

The functions  $s_i$  are step functions, with  $s_i(t)$  equal to one for  $\sum_{j=1}^{i-1} \tau_j \leq t < \sum_{j=1}^i \tau_j$  and equal to zero elsewhere, where the durations  $\tau_1, \dots, \tau_B$  are independent (not necessarily identically distributed) random variables with unknown distribution with support in  $[0, \tau_{\max}]$  and such that  $\bar{\tau}_1 \leq \mathbb{E}[\tau_i] \leq \bar{\tau}_2$ . Observing that the average duration of the earthquake is  $\sum_{i=1}^B \mathbb{E}[\tau_i]$  and keeping in mind the significant effect of this duration on structural reliability [97], we select  $\bar{\tau}_1 = 1 \text{ s}$ ,  $\bar{\tau}_2 = 2 \text{ s}$ ,  $\tau_{\max} = 6 \text{ s}$ , and  $B = 20$ .

The propagation through the earth structure gives rise to focusing, de-focusing, reflection, refraction and anelastic attenuation (which is caused by the conversion of wave energy to heat) [87]. We do not assume the earth structure to be known, henceforth we assume that  $\psi$  is a random transfer function of unknown distribution. More precisely, we assume that the transfer function is given by

$$\psi(t) := \frac{\sqrt{q}}{\tau'} \sum_{i=1}^q c_i \varphi_i(t), \quad (7.12)$$

where  $q := 20$ ,  $\tau' = 10 \text{ s}$ ,  $c$  is a random vector of unknown distribution with support in  $\{x \in [-1, 1]^q \mid \sum_{i=1}^q x_i^2 \leq 1 \text{ and } \sum_{i=1}^q x_i = 0\}$  and  $\varphi_i$  is a piecewise linear basis nodal element on the discretization  $t_1, \dots, t_q$  of  $[-\tau'/2, \tau'/2]$  with  $t_{i+1} - t_i = \tau'/q$  ( $\varphi_i(t_j) = \delta_{ij}$ , with  $\delta_{ij} = 1$  if  $i = j$  and zero otherwise).  $\psi$  has the dimension of 1/time and the constraint  $\sum_{i=1}^q c_i^2 \leq 1$  is equivalent to the assumption that  $(\frac{1}{\tau'} \int_{-\tau'/2}^{\tau'/2} |\psi|^2(t) dt)^{\frac{1}{2}}$  is, with probability one, bounded by a constant of order  $1/\tau'$ . Analogously to the Green function of the wave operator,  $\psi$  can take both positive and negative values (in time, for a fixed site and source). Observe also that the constraint on the time integral of  $\psi^2$  leads to a bound on the Arias intensity (i.e., the time integral of  $(\ddot{u}_0)^2$ ), which is a popular measure of ground motion strength used as a predictor of the likelihood of damage to short-period structures [86]. The constraint  $\sum_{i=1}^q c_i = 0$  ensures that the residual velocity is zero at the end of the earthquake. Observe also that, since the maximum amplitude of  $s$  already contains the dampening factor associated with the distance  $R$  to the center of the earthquake (in  $1/(R_0 + R)^2$ , via (7.11)),  $\psi$  has to be interpreted as a normalized transfer function. Since propagation in anisotropic structures can lead to changes in the direction of displacements, the coefficients  $c_i$  should, for full generality, be assumed to be tensors. Although we have assumed those coefficients to be scalars for the clarity and conciseness of the presentation, the method and reduction theorems proposed in this paper still apply when those coefficients are tensors.

### 7.1.3 The OUQ optimization problem

The optimal bound on the probability that the structure will fail is therefore the solution of the following optimization problem

$$\mathcal{U}(\mathcal{A}) := \sup_{(F, \mu) \in \mathcal{A}} \mu[F \leq 0], \quad (7.13)$$

where  $\mathcal{A}$  is the set of pairs  $(F, \mu)$  such that **(1)**  $F$  is mapping of the ground acceleration  $t \mapsto \ddot{u}_0(t)$  onto the margin  $\min_{i=1, \dots, J} (S_i - \|Y_i\|_\infty)$  via equations (7.8), (7.6) and (7.5). **(2)**  $\mu$  is a probability measure on the ground acceleration  $t \mapsto \ddot{u}_0(t)$  with support on accelerations defined by (7.9), (7.10), (7.12) (with  $B = 20$ ). Under this measure,  $X_1, \dots, X_B$ ,  $\tau_1, \dots, \tau_B$ ,  $c$  are independent (not necessarily identically distributed) random variables. For  $i = 1, \dots, B$ ,  $X_i$  has zero mean and independent (not necessarily identically distributed) components  $(X_{i,1}, X_{i,2}, X_{i,3})$  with support in  $[-a_{\max}, a_{\max}]$ , the measure of  $\tau_i$  is constrained by  $\bar{\tau}_1 \leq \mathbb{E}[\tau_i] \leq \bar{\tau}_2$  and has support in  $[0, \tau_{\max}]$ . The support of the measure on  $c$  is a subset of  $\{x \in [-1, 1]^q : \sum_{i=1}^q x_i^2 \leq 1 \text{ \& } \sum_{i=1}^q x_i = 0\}$ .

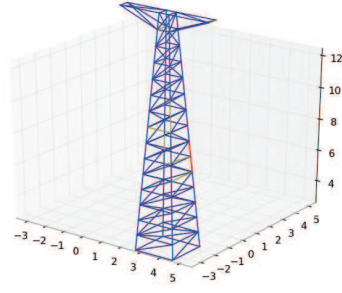
### 7.1.4 Reduction of the optimization problem

Problem (7.13) is not computationally tractable since the optimization variables take values in infinite-dimensional spaces of measures. However, thanks to Corollary 4.4, we know that the optimum of Problem (7.13) can be achieved by **(1)** Handling  $c$  as a deterministic optimization variable taking values in  $\{x \in [-1, 1]^q : \sum_{i=1}^q x_i^2 \leq 1 \text{ \& } \sum_{i=1}^q x_i = 0\}$  (since no constraints are given on the measure of  $c$ ) **(2)** Assuming that the measure on each  $X_{i,j}$  ( $X_i = (X_{i,1}, X_{i,2}, X_{i,3})$ ) is the tensorization of two Dirac masses in  $[-a_{\max}, a_{\max}]$  (since  $\mathbb{E}[X_{i,j}] = 0$  is one linear constraint) **(3)** Assuming that the measure on each  $\tau_i$  is the convex linear combination of 2 Dirac masses in  $[0, \tau_{\max}]$  ( $\bar{\tau}_1 \leq \mathbb{E}[\tau_i] \leq \bar{\tau}_2$  counts as one linear constraint).

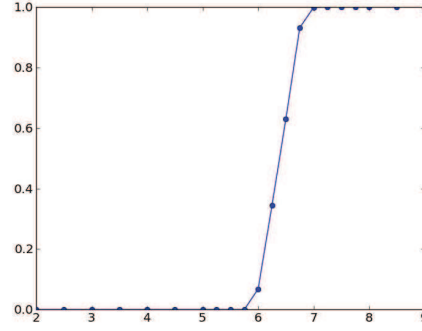
Observe that this reduced problem is of finite dimension ( $8B + q = 180$ ) (counting the scalar position of the Dirac masses, their weights and subtracting the number of scalar equality constraints).

### 7.1.5 Numerical results

The truss structure is the electrical tower shown in Sub-figure 7.1(a). This structure has 198 elements and we refer to [84] for precise numerical values associated with its geometry. The material used for this structure is steel. The corresponding material properties are  $7860 \text{ kg/m}^3$  for density,  $2.1 \cdot 10^{11} \text{ N/m}^2$  for the Young's modulus,  $2.5 \cdot 10^8 \text{ N/m}^2$  for the yield stress and  $\zeta = 0.07$  for the (uniform) damping ratio. Calculations were performed with time-step  $\Delta t := 5.0 \cdot 10^{-2} \text{ s}$ . We refer to Sub-figure 7.2(a) for a graph of the optimal bound on the probability of failure (7.13) versus the maximum ground acceleration 7.11 (in  $\text{m} \cdot \text{s}^{-2}$ ). Using Esteva's semi empirical formula (7.11) with a hypocentral distance  $R$  equal to 25 km we obtain Sub-figure 7.1(b), the graph of the optimal bound on the probability of failure (7.13) versus the earthquake of magnitude  $M_L$  in the Richter (local



(a) The truss structure



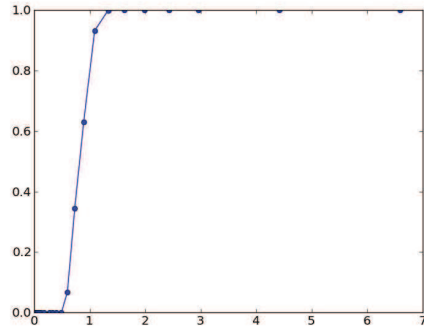
(b) Maximum PoF vs  $M_L$

Figure 7.1: Numerical results associated with the information set defined in Sub-section 7.1.2.

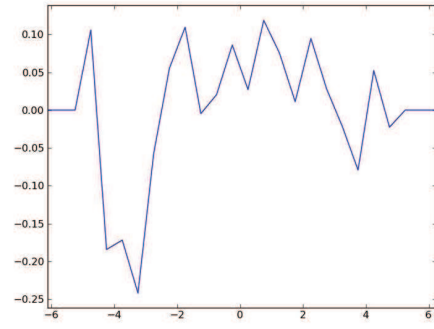
magnitude) scale at hypocentral distance  $R$  (the difference  $\Delta M_L$  between two consecutive points is 0.25). The “S” shape of the graph is typical of vulnerability curves [51]. We select one of the points in the transition region for further analysis — the point corresponding to a probability of failure of 0.631, a maximum ground acceleration of  $0.892 \text{ m} \cdot \text{s}^{-2}$  and an earthquake of magnitude 6.5. The vulnerability curve undergoes a sharp transition (from small probabilities of failures to unitary probabilities of failures) around maximum ground acceleration  $a_{\max} = 0.892 \text{ m} \cdot \text{s}^{-2}$ . This transition becomes smoother as the number of independent variables in the description of the admissible set is increased (results not shown).

For  $a_{\max} = 0.892 \text{ m} \cdot \text{s}^{-2}$  ( $M_L = 6.5$ ), Sub-figures 7.2(b) and 7.2(c) show the (deterministic) transfer function  $\psi$  (the units in the x-axis are seconds) and 3 independent realizations of the earthquake source  $s(t)$  sampled from the measure  $\mu_{0.892}$  maximizing the probability of failure. For this measure, Sub-figure 7.2(f) shows the axial strain of all elements versus time (in seconds) and Sub-figure 7.1(a) identifies the ten weakest elements for the most probable earthquake (the axial strain of these elements are: 0.00142317, 0.00125928, 0.00099657, 0.00081897, 0.00076223, 0.00075958, 0.00072190, 0.00068266, 0.00062919, and 0.00061361) — the weakest two elements exceed the yield strain of 0.00119048 (shown in red in the figure). Sub-figures 7.2(d) and 7.2(e) show 3 independent horizontal ground acceleration and a power spectrum sampled from  $\mu_{0.892}$ . The units in Sub-figure 7.2(e) are cycles per seconds for the  $x$  axis and  $\text{m} \cdot \text{s}^{-2}$  for the  $y$  axis. The units in Sub-figure 7.2(d) are seconds for the  $x$  axis and  $\text{m} \cdot \text{s}^{-2}$  for the  $y$  axis.

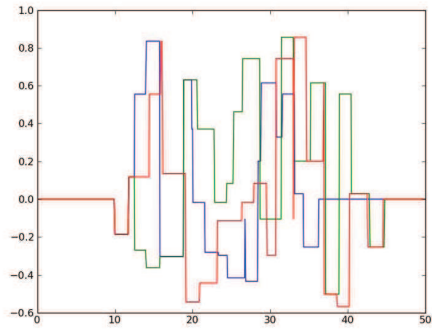
An quantitative analysis of the numerical results also show that all the constraints are active at the extremum (i.e. the generalized moments inequality constraints on  $\mu$  defining the information set introduced in Sub-section 7.1.2 are equalities or near equalities at the extremum) . The positions and weights of the Dirac masses associated



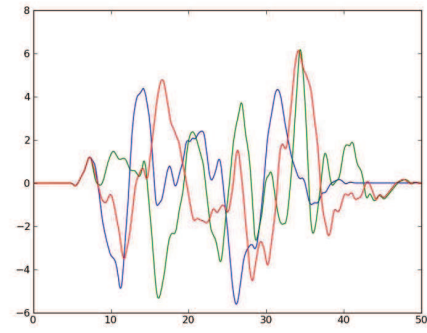
(a) Maximum PoF vs  $a_{max}$



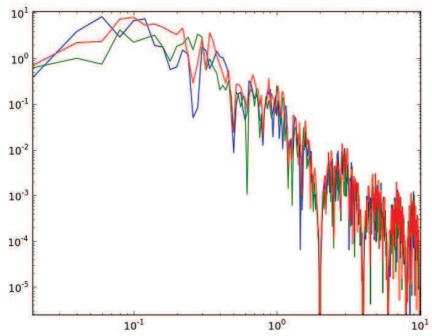
(b) Transfer function  $\psi$



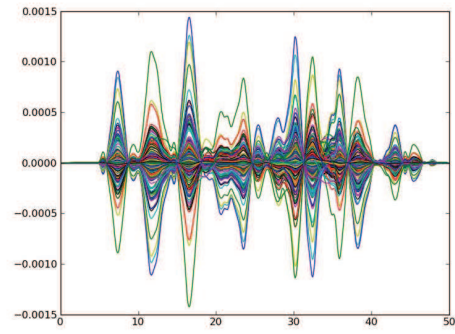
(c) Earthquake source  $s(t)$



(d) Ground acceleration



(e) Power Spectrum



(f) Elements strain

Figure 7.2: Numerical results associated with the information set defined in Sub-section 7.1.2.

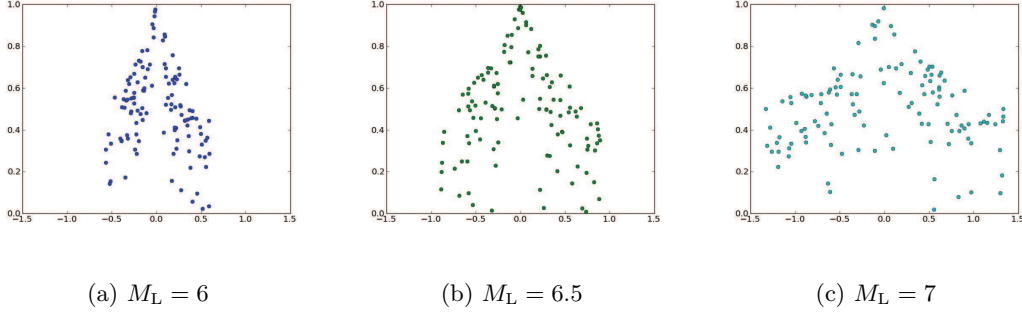


Figure 7.3: Positions (abscissa, in  $\text{m} \cdot \text{s}^{-2}$ ) and weights (ordinates) of the Dirac masses associated with the measure of probability on  $X_1, \dots, X_B$  at the extremum for earthquakes of magnitude  $M_L = 6$ ,  $M_L = 6.5$  and  $M_L = 7$ . Note that the positions in abscissa correspond to the possible amplitudes of the impulses  $X_i$ .

with durations and transfer coefficients do not appear to show any discernible trend. However, the positions and weights of the Dirac masses associated with the amplitudes  $X_1, \dots, X_M$  show a trend (as function of the earthquake magnitude  $M_L$ ) illustrated in Figure 7.3. This trend suggests that for strong earthquakes, probabilities of failures are maximized via (the possibility of) large amplitude impulses.

**On the numerical optimization algorithm.** Global search algorithms often require hundreds of iterations and thousands of function evaluations, due to their stochastic nature, to find a global optimum. Local methods, like Powell’s method [71], may require orders of magnitude fewer iterations and evaluations, but do not generally converge to a global optimum in a complex parameter space. To compute the probability of failure, we use a Differential Evolution algorithm [72, 88] that has been modified to utilize large-scale parallel computing resources [59]. Each iteration, the optimizer prepares  $m$  points in parameter space, with each new point derived through random mutations from the ‘best’ point in the previous iteration. We select  $m = 40$ , which is of modest size compared to the dimensionality of the problem — however, we chose this modest size because populations larger than  $m = 40$  only modestly increase the efficiency of the algorithm. Each of these  $m$  evaluations are performed in parallel on a computer cluster, such that the time required for a single iteration equals the time required for a single function evaluation. After  $n$  iterations complete, the optimal probability of failure for the product measure is returned (convergence is observed around  $n \approx 200$  and we select  $n \approx 2000$  for the robustness of the result).

Only one iteration is required for values of ground acceleration on the extremes of the range (such as  $M_L = 2$  and  $M_L = 9$ ). The number of iterations required for convergence for points in the transition region (around  $M_L = 6.5$ ) is between 30 and 50 (which corresponds to 2,400 to 4,000 function evaluations). We refer to Figure 7.4 for

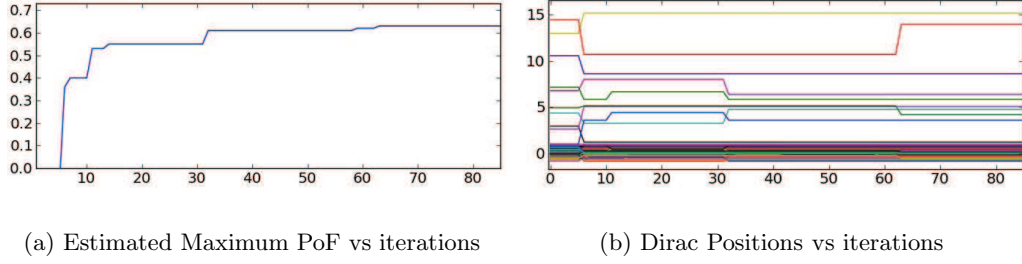


Figure 7.4: (a): Estimated maximum probability of failure versus number of iterations for an earthquake of magnitude  $M_L = 6.5$  (this corresponds to the point in transition region of Sub-figure 7.1(b)). (b): re-normalized positions of the Dirac masses for  $M_L = 6.5$ .

an illustration of the convergence of the optimization algorithm for  $M_L = 6.5$ .

Each function evaluation takes approximately 0.5 s on a high-performance computing cluster (such as the high-performance computing clusters used at the National Labs). With each iteration utilizing  $m = 40$  parallel processors, the OUQ calculation takes roughly 24 hrs.

Approximately 1000 time steps are required for accuracy in the strain calculations, each function evaluation requires two convolutions over time. Because of the size of the truss structure (198 elements), eigenvalues have to be computed with high accuracy. Because of the size of the product measure associated with the numerical optimization iterates, the probability of failure (associated with these iterates) should be estimated with a controlled (and adapted) tolerance rather than computed exactly — we use a sampling size of 5000 points.

## 7.2 OUQ and critical excitation.

Without constraints on ground acceleration, the ground motion yielding the maximum peak response (maximum damage in a deterministic setting) has been referred to as the *critical excitation* [20]. Drenick himself pointed out that a seismic design based on critical excitation could be “far too pessimistic to be practical” [21]. He later suggested that the combination of probabilistic approaches with worst-case analysis should be employed to make the seismic resistant design robust [22]. Practical application and extension of critical excitation methods have then been made extensively and we refer to [92] and [93] for recent reviews. The probabilities of failures obtained from stochastic approaches depend on particular choices of probability distribution functions. Because of the scarcity of recorded time-histories, these choices involve some degree of arbitrariness [87, 92] that may be incompatible with the certification of critical structures and rare events [23]. We suggest that by allowing for very weak assumption on probability measures, the reduction theorems associated with the OUQ framework could lead to certifications methods that are both robust (reliable) and practical (not overly pessimistic). Of course this does

require the identification of a reliable and narrow information set. The set  $\mathcal{A}$  used in this paper does not include all the available information on earthquakes. We also suggest that the method of selecting next best experiments could help in this endeavor.

Observe also that without constraints, worst-case scenarios correspond to focusing the energy of the earthquake in modes of resonances of the structure. Without correlations in the ground motion these scenarios correspond to rare events where independent random variables must conspire to strongly excite a specific resonance mode. The lack of information on the transfer function  $\psi$  and the mean values  $\mathbb{E}[\tau_i]$  permits scenarios characterized by strong correlations in ground motion where the energy of the earthquake can be focused in the above mentioned modes of resonance.

### 7.3 Alternative formulation in the frequency domain

A popular method for modeling and synthesizing seismic ground motion is to use (deterministic) shape functions and envelopes in the frequency domain (see [96] for a review).

In this sub-section, we will evaluate the safety of the electrical tower shown in Sub-figure 7.1(a) using an admissible set  $\mathcal{A}_F$  defined from weak information on the probability distribution of the power spectrum of the seismic ground motion.

#### 7.3.1 Formulation of the information set

We assume that the (three dimensional) ground motion acceleration is given by

$$\ddot{u}_0(t) := \sum_{k=1}^W ((A_{6k-5}, A_{6k-4}, A_{6k-3}) \cos(2\pi\omega_k t) + (A_{6k-2}, A_{6k-1}, A_{6k}) \sin(2\pi\omega_k t)), \quad (7.14)$$

where the Fourier coefficients  $A_j$  are random variables (in  $\mathbb{R}$ ) of unknown distribution. We assume that  $W := 100$  and that  $\omega_k := k/\tau_d$  with  $\tau_d = 20$  s. Writing  $A := (A_1, \dots, A_{6W})$ , we assume that

$$\mathbb{P}[A \in B(0, a_{\max}) \setminus B(0, \frac{a_{\max}}{2})] = 1, \quad (7.15)$$

where  $a_{\max}$  is given by Esteve's semi-empirical expression (7.11) and  $B(0, a_{\max}) \setminus B(0, \frac{a_{\max}}{2})$  is the Euclidean ball of  $\mathbb{R}^{6W}$  of center 0 and radius  $a_{\max}$  minus the Euclidean ball of center 0 and radius  $\frac{a_{\max}}{2}$ .

Although different earthquakes have different power spectral densities it is empirically observed that “on average”, their power spectra follow specific shape functions that may depend on the ground structure of the site where the earthquake is occurring [47]. Based on this observation, synthetic seismograms are produced by filtering the Fourier spectrum of white noise with these specific shape functions [47]. In this sub-section, our information on the distribution of  $A$  will be limited to the shape of the mean value of its power spectrum. More precisely, we will assume that, for  $k \in \{1, \dots, W\}$  and  $j \in \{0, \dots, 5\}$ ,

$$\mathbb{E}[A_{6k-j}^2] = \frac{a_{\max}^2}{12} \frac{s(\omega_k)}{s_0}, \quad (7.16)$$



where  $s$  is the Matsuda–Asano shape function [56] given by:

$$s(\omega) := \frac{\omega_g^2 \omega^2}{(\omega_g^2 - \omega^2)^2 + 4\xi_g^2 \omega_g^2 \omega^2}, \quad (7.17)$$

where  $\omega_g$  and  $\xi_g$  are the natural frequency and natural damping factor of the site and

$$s_0 := \sum_{k=1}^W s(\omega_k). \quad (7.18)$$

We will use the numerical values  $\omega_g = 6.24$  Hz and  $\xi_g = 0.662$  associated with the January 24, 1980 Livermore earthquake (see [49], observe that we are measuring frequency in cycles per seconds instead of radians per seconds). The purpose of the normalization factor (7.18) is to enforce the following mean constraint:

$$\mathbb{E} \left[ \frac{1}{\tau_d} \int_0^{\tau_d} |\ddot{u}_0(t)|^2 dt \right] = \frac{1}{2} \mathbb{E} [|A|^2] = \frac{a_{\max}^2}{4}. \quad (7.19)$$

Observe also that (7.15) implies that, with probability one,

$$\frac{a_{\max}^2}{8} \leq \frac{1}{\tau_d} \int_0^{\tau_d} |\ddot{u}_0(t)|^2 dt \leq \frac{a_{\max}^2}{2}. \quad (7.20)$$

We write  $\mathcal{A}_F$  the set of probability measures  $\mu$  on  $A$  satisfying (7.15) and (7.16).

### 7.3.2 OUQ objectives

Let  $(Y_1, \dots, Y_J)$  and  $(S_1, \dots, S_J)$  be the axial and yield strains introduced in Sub-section 7.1.1. Writing  $S := [-S_1, S_1] \times \dots \times [-S_J, S_J]$  (this is the safe domain for the axial strains), we are interested in computing optimal (maximal and minimal with respect to measures  $\mu \in \mathcal{A}_F$ ) bounds on the probability (under  $\mu$ ) that  $Y(t) \notin S$  for some  $t \in [0, \tau_d]$  (defined as the probability of failure). From the linearity of equations (7.3), the strain of member  $i$  ( $i \in \{1, \dots, J\}$ ) at time  $t$  can be written

$$Y_i(t) = \sum_{j=1}^{6W} \Psi_{ij}(t) A_j. \quad (7.21)$$

Let  $\Psi(t)$  be the  $J \times (6W)$  tensor  $(\Psi_{ij}(t))$  and observe that Equation (7.21) can be also be written  $Y(t) = \Psi(t)A$ . Let  $\mathcal{F}$  be the subset of  $\mathbb{R}^{6W}$  defined as the elements  $x$  of  $B(0, a_{\max}) \setminus B(0, \frac{a_{\max}}{2})$  such that  $\Psi(t)x \notin [-S_1, S_1] \times \dots \times [-S_J, S_J]$  for some  $t \in [0, \tau_d]$ , i.e.

$$\mathcal{F} := \left\{ x \in B(0, a_{\max}) \setminus B(0, \frac{a_{\max}}{2}) \mid \Psi(t)x \notin S \text{ for some } t \in [0, \tau_d] \right\}. \quad (7.22)$$

Observe  $\mathcal{F}$  corresponds to the set of vectors  $A$  (in (7.14)) that lead to a failure of the structure. Henceforth, our objective can be formulated as computing

$$\sup_{\mu \in \mathcal{A}_F} \mu[A \in \mathcal{F}] \quad \text{and} \quad \inf_{\mu \in \mathcal{A}_F} \mu[A \in \mathcal{F}], \quad (7.23)$$



where  $\mathcal{A}_F$  is the set of probability measures  $\mu$  such that  $\mu[A \in B(0, a_{\max}) \setminus B(0, \frac{a_{\max}}{2})] = 1$ , and that

$$\mathbb{E}_\mu[A_j^2] = b_j \quad \text{with} \quad b_j := \frac{a_{\max}^2}{12} \frac{s(\omega_{\lfloor (j+5)/6 \rfloor})}{s_0}. \quad (7.24)$$

In other words,  $\mathcal{A}_F$  an infinite-dimensional polytope defined as the set of probability measures on ground acceleration that have the Matsuda–Asano average power spectra (7.17). It is important to observe that with the filtered white noise method the safety of the structure is assessed for a single measure  $\mu_0 \in \mathcal{A}_F$  whereas in the proposed OUQ framework we compute best and worst-case scenarios with respect to all measures in  $\mathcal{A}_F$ .

### 7.3.3 Reduction of the optimization problem with Dirac masses

Since (7.24) corresponds to  $6W$  global linear constraints on  $\mu$ , Theorem 4.1 implies that the extrema of problem (7.23) can be achieved by assuming  $\mu$  to be a weighted sum of Dirac masses  $\sum_{j=1}^{6W+1} p_j \delta_{Z_{\cdot,j}}$  where  $Z_{\cdot,j} \in B(0, a_{\max}) \setminus B(0, \frac{a_{\max}}{2})$ ,  $p_j \geq 0$  and  $\sum_{j=1}^{6W+1} p_j = 1$ . The constraints (7.24) can then be written: for  $i \in \{1, \dots, 6W\}$ ,  $\sum_{j=1}^{6W+1} Z_{i,j}^2 p_j = b_i$ . Furthermore,  $\mu[A \in \mathcal{F}] = \sum_{j: Z_{\cdot,j} \in \mathcal{F}} p_j$ .

### 7.3.4 Reduction of the optimization problem based on strong duality

Since the information contained in  $\mathcal{A}_F$  is limited to constraints on the moments of  $A$ , strong duality can be employed to obtain an alternative reduction of problems (7.23). Indeed, Theorem 2.2 of [13] implies that

$$\sup_{\mu \in \mathcal{A}_F} \mu[A \in \mathcal{F}] = \inf_{(H_0, H) \in \mathbb{R}^{6W+1}} H_0 + \sum_{i=1}^{6W} H_i b_i, \quad (7.25)$$

where the minimization problem (over the vector  $(H_0, H) := (H_0, H_1, \dots, H_{6W}) \in \mathbb{R}^{6W+1}$ ) in the right hand side of (7.25) is subject to

$$\sum_{i=1}^{6W} H_i x_i^2 + H_0 \geq \chi(x) \quad \text{on} \quad B(0, a_{\max}) \setminus B(0, \frac{a_{\max}}{2}), \quad (7.26)$$

where  $\chi(x)$  is the function equal to 1 on  $\mathcal{F}$  and 0 on  $(\mathcal{F})^c$  (we note  $(\mathcal{F})^c$  the complement of  $\mathcal{F}$ , i.e. the set of  $x$  in  $\mathbb{R}^{6W}$  that are not elements of  $\mathcal{F}$ ). Similarly,

$$\inf_{\mu \in \mathcal{A}_F} \mu[A \in \mathcal{F}] = \sup_{(H_0, H) \in \mathbb{R}^{6W+1}} H_0 + \sum_{i=1}^{6W} H_i b_i, \quad (7.27)$$

where the maximization problem in the right hand side of (7.27) is subject to

$$\sum_{i=1}^{6W} H_i x_i^2 + H_0 \leq \chi(x) \quad \text{on} \quad B(0, a_{\max}) \setminus B(0, \frac{a_{\max}}{2}). \quad (7.28)$$

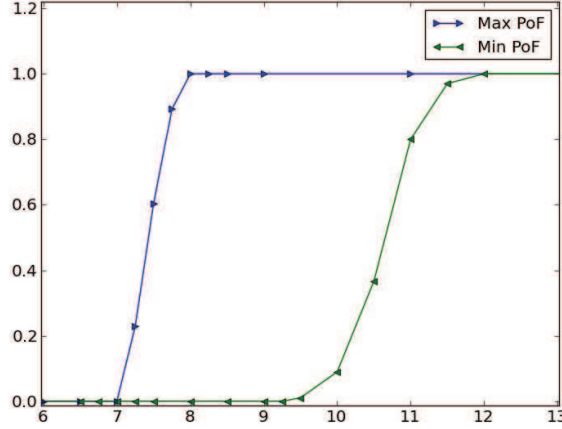


Figure 7.5: Maximum and minimum probability of failure of the structure (as defined in (7.23)) versus the earthquake of magnitude  $M_L$  in the Richter (local magnitude) scale at hypocentral distance  $R = 25$  km ( $a_{\max}$  is given by Esteva's semi-empirical expression (7.11) as a function of  $M_L$ ). The curve corresponding to maximum probability of failure is not the same as the one given in Sub-figure 7.1(b) because it is based on a different information set.

We conclude from these equations (by optimizing first with respect to  $H_0$ ) that the optimal upper bound on the probability of failure (defined as the probability that the displacement  $Y(t)$  does not belong to the safe region  $S$  for all time  $t$  in the interval  $[0, \tau_d]$ ) is

$$\sup_{\mu \in \mathcal{A}_F} \mu[A \in \mathcal{F}] = \inf_{H \in \mathbb{R}^{6W}} \sup_{x \in B(0, a_{\max})/B(0, \frac{a_{\max}}{2})} \chi(x) + \sum_{i=1}^{6W} H_i(b_i - x_i^2), \quad (7.29)$$

whereas the optimal lower bound is

$$\inf_{\mu \in \mathcal{A}_F} \mu[A \in \mathcal{F}] = \sup_{H \in \mathbb{R}^{6W}} \inf_{x \in B(0, a_{\max})/B(0, \frac{a_{\max}}{2})} \chi(x) + \sum_{i=1}^{6W} H_i(b_i - x_i^2). \quad (7.30)$$

Observe that problem (7.29) is convex in  $H \in \mathbb{R}^{6W}$  whereas problem (7.30) is concave.

### 7.3.5 Numerical results

The optimal bounds (7.23) can be computed using the reduction to masses to Dirac described in Sub-section 7.3.3 or strong duality as described in Sub-section 7.3.4. While the latter does not identify the extremal measures it leads to a smaller optimization problem than the former (i.e. to optimization variables in  $\mathbb{R}^{12W}$ , instead of  $\mathbb{R}^{(6W+1) \times (6W+1)}$ ).

The simplification is allowed by the facts that the response function is well identified, that there are no independence constraints, and that the information on  $A$  is limited to  $6W$  (scalar) moment constraints. The vulnerability curves of Figure 7.5 have been computed using strong duality as described in Subsection 7.3.4 (an identification of extremal measures would require using the method described in Subsection 7.3.3). Observe that to decrease the gap between the maximum probability of failure and the minimum probability of failure, one would have to refine the information on the probability distribution of ground motion (by, for instance, adding constraints involving the correlation between the amplitudes  $A_i$  different Fourier modes). To solve optimization problems (7.29) and (7.30) we use the modified Differential Evolution algorithm described in Subsection 7.1.5. Equation (7.29) is implemented as a minimization over  $H$ , where a nested maximization over  $x$  is used to solve for  $\sup_{x \in B(0, a_{\max})/B(0, \frac{a_{\max}}{2})} \chi(x) - \sum_{i=1}^{6W} H_i x_i^2$  at each function evaluation. Both the minimization over  $H$ , and the maximization over  $x$  use the Differential Evolution algorithm described above, where the optimizer configuration itself differs only in that for the nested optimization termination occurs when the maximization over  $x$  does not improve by more than  $10^{-6}$  in 20 iterations, while the outer optimization is terminated when there is not more than  $10^{-6}$  improvement over 100 iterations. The optimization over  $H$  is performed in parallel, as described in Subsection 7.1.5, where each of the nested optimizations over  $x$  are distributed across nodes of a high-performance computing cluster. Each of the (nested) optimizations over  $x$  require only a few seconds on average, and thus are performed serially. Convergence, on average takes about 15 hours, and is obtained in roughly 2000 iterations (over  $H$ ), corresponding to 35000 to 50000 function evaluations. Each function evaluation is a nested optimization over  $x$ , which takes a few seconds on a high-performance computing cluster.

## 8 Application to Transport in Porous Media

We now apply the OUQ framework and reduction theorems to divergence form elliptic PDEs and consider the situation where coefficients (corresponding to microstructure and source terms) are random and have imperfectly known probability distributions. Treating those distributions as optimization variables (in an infinite-dimensional space) we obtain optimal bounds on probabilities of deviation of solutions. Surprisingly, explicit and optimal bounds show that, with incomplete information on the probability distribution of the microstructure, uncertainties or information do not necessarily propagate across scales.

To make this more precise in a simple setting, let  $\mathcal{D} \subseteq \mathbb{R}$  be a bounded domain and consider  $u(x, \omega)$ , the solution of the following stochastic elliptic PDE:

$$\begin{cases} -\operatorname{div}(\kappa(x, \omega) \nabla u(x, \omega)) = f(x, \omega), & x \in \mathcal{D} \\ u(x, \omega) = 0, & x \in \partial\mathcal{D} \end{cases} \quad (8.1)$$

with random microstructure  $\kappa$  and random (positive) source term  $f$ . Physically,  $u$  can be interpreted as the pressure (or head) in a medium of permeability  $\kappa$  with source  $f$ ; the

fluid velocity is given by  $\nabla u$ . For a given point  $x_0$  in the interior of  $\mathcal{D}$ , we are interested in computing the least upper bound on the probability of an unsafe supercritical pressure at  $x_0$ :

$$\mathcal{U}(\mathcal{A}) := \sup_{\mu \in \mathcal{A}} \mathbb{E}_\mu [\log u(x_0, \omega) \geq \mathbb{E}_\mu [\log u(x_0, \omega)] + a], \quad (8.2)$$

where  $\mathcal{A}$  is a set of probability measures on  $(\kappa, f)$ . In this section we will focus on the two admissible sets  $\mathcal{A}$  described below.

Let  $D_1, D_2 \geq 0$ ,  $K, F \in L^\infty(\mathcal{D})$  such that  $\text{essinf}_{\mathcal{D}} K > 0$ ,  $F \geq 0$ , and  $\int_{\mathcal{D}} F > 0$ . Define

$$\mathcal{A}_{\kappa, f} := \left\{ \mu \left| \begin{array}{l} \kappa, f \text{ independent under } \mu, \\ K(x) \leq \kappa(x, \omega) \leq e^{D_1} K(x), \\ F(x) \leq f(x, \omega) \leq e^{D_2} F(x) \end{array} \right. \right\}. \quad (8.3)$$

We say that a function  $g$  defined on  $\mathcal{D}$  is periodic of period  $\delta$  if for all  $x \in \mathcal{D}$ , it holds that  $g(x) = g(x + \delta)$  whenever  $x + \delta \in \mathcal{D}$ . We now define

$$\mathcal{A}_{\kappa_1 \kappa_2} := \left\{ \mu \left| \begin{array}{l} \kappa = \kappa_1 \kappa_2, \\ \kappa_1, \kappa_2 \text{ independent under } \mu, \\ \|\nabla \kappa_1\|_{L^\infty} \leq e^{D_1} \|\nabla K_1\|_{L^\infty} \\ \kappa_2 \text{ periodic of period } \delta \\ K_1(x) \leq \kappa_1(x, \omega) \leq e^{D_1} K_1(x), \\ K_2(x) \leq \kappa_2(x, \omega) \leq e^{D_2} K_2(x), \end{array} \right. \right\}, \quad (8.4)$$

where  $0 < \delta \ll 1$ ,  $K_2 \in L^\infty(\mathcal{D})$  is uniformly elliptic over  $\mathcal{D}$  and periodic of period  $\delta$ , and  $K_1$  is smooth and uniformly elliptic over  $\mathcal{D}$ .

PDEs of the form (8.1) have become a benchmark for stochastic expansion methods [29, 101, 4, 26, 19, 94, 14] and we also refer to [30] for their importance for transport in porous media.

These PDEs have also been studied as classical examples in the UQ literature on the basis that the randomness in the coefficients (with a perfectly known probability distribution on the coefficients  $(\kappa, f)$ ) is an adequate model of the lack of information on the microstructure  $\kappa$ . In these situations the quantification of uncertainties is equivalent to a push forward of the measure probability on  $(\kappa, f)$ .

However, in practical situations the probability distribution on the coefficients  $(\kappa, f)$  may not be known a priori and the sole randomness in coefficients may not constitute a complete characterization of uncertainties. This is our motivation for considering the problem described in this section. We have also introduced the admissible set (8.4) as a simple illustration of uncertainty quantification with multiple scales and incomplete information on probability distributions. To relate this example to classical homogenization [8] we have assumed  $\kappa_2$  to be periodic of small period  $\delta \ll 1$ .

**Theorem 8.1.** *We have*

$$\mathcal{U}(\mathcal{A}_{\kappa, f}) = \mathcal{U}(\mathcal{A}_{\kappa_1 \kappa_2}) = \mathcal{U}(\mathcal{A}_{\text{McD}}), \quad (8.5)$$

with

$$\mathcal{U}(\mathcal{A}_{\text{McD}}) = \begin{cases} 0, & \text{if } D_1 + D_2 \leq a, \\ \frac{(D_1 + D_2 - a)^2}{4D_1D_2}, & \text{if } |D_1 - D_2| \leq a \leq D_1 + D_2, \\ 1 - \frac{a}{\max(D_1, D_2)}, & \text{if } 0 \leq a \leq |D_1 - D_2|. \end{cases} \quad (8.6)$$

Before giving the proof of Theorem 8.1, we make a few important observations:

It follows from Theorem 8.1 that if  $D_2 \geq a + D_1$ , then  $\mathcal{U}(\mathcal{A}_{\kappa,f})(a, D_1, D_2) = \mathcal{U}(\mathcal{A}_{\kappa,f})(a, 0, D_2)$ . In other words, if the uncertainty on the source term  $f$  is dominant, then the uncertainty associated with the microstructure,  $\kappa$ , does not propagate to the uncertainty corresponding to the probability of deviation of  $\log u(x_0, \omega)$  from its mean.

Now consider  $\mathcal{A}_{\kappa_1\kappa_2}$ . Since  $\kappa_1$  is constrained to be smooth and  $\kappa_2$  periodic with period  $\delta \ll 1$ , one would expect the microstructure  $\kappa_2$  to appear in the probability of deviation in a homogenized form. However, Theorem 8.1 shows that if  $D_1 \geq a + D_2$ , then  $\mathcal{U}(\mathcal{A}_{\kappa_1\kappa_2})(a, D_1, D_2) = \mathcal{U}(\mathcal{A}_{\kappa_1\kappa_2})(a, D_1, 0)$ . More precisely, if the uncertainty associated with the background  $\kappa_1$  is dominant, then the uncertainty associated with the microstructure  $\kappa_2$  does not propagate to the uncertainty corresponding to the probability of deviation of  $\log u(x_0, \omega)$  from its mean.

This simple but generic example suggests that for structures characterized by multiple scales or multiple modules, information or uncertainties may not propagate across modules or scales. This phenomenon can be explained by the fact that, with incomplete information, scales or modules may not communicate certain types of information. Henceforth, the global uncertainty of a modular system cannot be reduced without decreasing local dominant uncertainties. In particular, for modular or multi-scale systems, one can identify (possibly large) accuracy thresholds (in terms of numerical solutions of PDEs or SPDEs) below which the global uncertainty of the system does not decrease.

*Proof of Theorem 8.1.* Let us now prove Theorem 8.1 with the admissible set  $\mathcal{A}_{\kappa,f}$  (the proof with the set  $\mathcal{A}_{\kappa_1\kappa_2}$  is similar). It follows from Theorem 2.11 and Proposition 2.13 of [66] that the maximum oscillation of  $\log u(x_0, \omega)$  with respect to  $\kappa$  and  $f$  are bounded by  $D_1$  and  $D_2$  we obtain that

$$\mathcal{U}(\mathcal{A}_{\kappa,f}) \leq \mathcal{U}(\mathcal{A}_{\text{McD}}), \quad (8.7)$$

where  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  is defined in equation (4.12) (we consider the case  $m = 2$ ).

Next, from the proof of Theorem 5.2, we observe that the bound  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  can be achieved by  $\mathcal{A}_{\kappa,f}$  by considering measures  $\mu$  that are tensorizations of two weighted Dirac masses in  $\kappa$  (placed at  $K$  and  $e^{D_1}K$ ) and two weighted Dirac masses in  $f$  (placed at  $F$  and  $e^{D_1}F$ ). This concludes the proof.  $\square$

## 9 Conclusions

**The UQ Problem — A Problem with UQ?** The 2003 *Columbia* space shuttle accident and the 2010 Icelandic volcanic ash cloud crisis have demonstrated two sides of the

same problem: discarding information may lead to disaster, whereas over-conservative safety certification may result in unnecessary economic loss and supplier-client conflict. Furthermore, while everyone agrees that UQ is a fundamental component of objective science (because, for instance, objective assertions of the validity of a model or the certification of a physical system require UQ), it appears that not only is there no universally accepted notion of the objectives of UQ, there is also no universally accepted framework for the communication of UQ results. At present, the “UQ problem” appears to have all the symptoms of an ill-posed problem; at the very least, it lacks a coherent general presentation, much like the state of probability theory before its rigorous formulation by Kolmogorov in the 1930s.

- At present, UQ is an umbrella term that encompasses a large spectrum of methods: Bayesian methods, Latin hypercube sampling, polynomial chaos expansions, stochastic finite-element methods, Monte Carlo, *etc.* Most (if not all) of these methods are characterized by a list of assumptions required for their application or efficiency. For example, Monte Carlo methods require a large number of samples to estimate rare events; stochastic finite-element methods require the precise knowledge of probability density functions and some regularity (in terms of decays in spectrum) for their efficiency; and concentration-of-measure inequalities require uncorrelated (or weakly correlated) random variables.
- There is a disconnect between theoretical UQ methods and complex systems of importance requiring UQ in the sense that the assumptions of the methods do not match the assumption/information set of the application. This disconnect means that often a specific method adds inappropriate implicit or explicit assumptions (for instance, when the knowledge of probability density functions is required for polynomial chaos expansions, but is unavailable) and/or the repudiation of relevant information (for instance, the monotonicity of a response function in a given variable) that the method is not designed to incorporate.

**OUQ as an opening gambit.** OUQ is not the definitive answer to the UQ problem, but we hope that it will help to stimulate a discussion on the development of a rigorous and well-posed UQ framework analogous to that surrounding the development of probability theory. The reduction theorems of Section 4, the Optimal Concentration Inequalities and non-propagation of input uncertainties of Section 5, the possibility of the selection of optimal experiments described at the end of Section 2, and the numerical evidence of Section 6 that (singular, i.e. low-dimensional) optimizers are also attractors, suggest that such a discussion may lead to non-trivial and worthwhile questions and results at the interface of optimization theory, probability theory, computer science and statistics.

In particular, many questions and issues raised by the OUQ formulation remain to be investigated. A few of those questions and issues are as follows:

- Any (possibly numerical) method that finds admissible states  $(f, \mu) \in \mathcal{A}$  leads to rigorous lower bounds on  $\mathcal{U}(\mathcal{A})$ . It is known that duality techniques lead to upper bounds on  $(f, \mu) \in \mathcal{A}$  provided that the associated Lagrangians can be computed.

Are there interesting classes of problems for which those Lagrangians can rigorously be estimated or bounded from above?

- The reduction theorems of Section 4 are limited to linear constraints on probability distribution marginals and the introduction of sample data may lead to other situations of interest (for instance, relative-entropy type constraints).
- Although general in its range of application, the algorithmic framework introduced in Section 6 is still lacking general convergence theorems.
- The introduction of sample data appears to render the OUQ optimization problem even more complex. Can this optimization problem be made equivalent to applying the deterministic setting to an information set  $\mathcal{A}$  randomized by the sample data?
- In the presence of sample data, instead of doing theoretical analysis to describe the optimal statistical test, one formulation of the OUQ approach provides an optimization problem that must be solved to determine the test. Is this optimization problem reducible?

## Acknowledgements

The authors gratefully acknowledge portions of this work supported by the Department of Energy National Nuclear Security Administration under Award Number DE-FC52-08NA28613 through Caltech’s ASC/PSAAP Center for the Predictive Modeling and Simulation of High Energy Density Dynamic Response of Materials. Calculations for this paper were performed using the *mystic* optimization framework [59]. We thank the Caltech PSAAP Experimental Science Group — Marc Adams, Leslie Lamberson, Jonathan Mihaly, Laurence Bodelot, Justin Brown, Addis Kidane, Anna Pandolfi, Guruswami Ravichandran and Ares Rosakis — for Formula (1.5) and figures 1.2. We thank Sydney Garstang and Carmen Sirois for proofreading the manuscript. We thank Ilse Ipsen and four anonymous referees for detailed comments and suggestions.

## 10 Appendix: Proofs

### 10.1 Proofs for Section 4

*Proof of Theorem 4.1.* In this proof, we use  $(\mu_1, \dots, \mu_m)$  as a synonym for the product  $\mu_1 \otimes \dots \otimes \mu_m$ . For  $\mu = \bigotimes_{i=1}^m \mu_i \in \mathcal{M}^{\mathbf{G}}$ , consider the optimization problem

$$\begin{aligned} & \text{maximize: } \mathbb{E}_{(\mu'_1, \mu_2, \dots, \mu_m)}[r], \\ & \text{subject to: } \mu'_1 \in \mathcal{M}(\mathcal{X}_1), \\ & \qquad \qquad \mathbf{G}(\mu'_1, \mu_2, \dots, \mu_m) \leq 0. \end{aligned}$$

By Fubini’s Theorem,

$$\mathbb{E}_{(\mu'_1, \mu_2, \dots, \mu_m)}[r] = \mathbb{E}_{\mu'_1} [\mathbb{E}_{(\mu_2, \dots, \mu_m)}[r]],$$

where  $\mathbb{E}_{(\mu_2, \dots, \mu_m)}[r]$  is a Borel-measurable function on  $\mathcal{X}_1$  and, for  $j = 1, \dots, n$ , it holds that

$$\mathbb{E}_{(\mu'_1, \mu_2, \dots, \mu_m)}[g'_j] = \mathbb{E}_{\mu'_1} [\mathbb{E}_{(\mu_2, \dots, \mu_m)}[g'_j]],$$

where  $\mathbb{E}_{(\mu_2, \dots, \mu_m)}g'_j$  is a Borel-measurable function on  $\mathcal{X}_1$ . In the same way, we see that

$$\mathbb{E}_{(\mu'_1, \mu_2, \dots, \mu_m)}[g_j^1] = \mathbb{E}_{\mu'_1}[g_j^1],$$

and, for  $k = 2, \dots, m$  and  $j = 1, \dots, n_k$ , it holds that

$$\mathbb{E}_{(\mu'_1, \mu_2, \dots, \mu_m)}[g_j^k] = \mathbb{E}_{\mu_k}[g_j^k],$$

which are constant in  $\mu'_1$ .

Since each  $\mathcal{X}_i$  is Suslin, it follows that all the measures in  $\mathcal{M}(\mathcal{X}_i)$  are regular. Consequently, by [95, Theorem 11.1], the extreme set of  $\mathcal{M}(\mathcal{X}_i)$  is the set of Dirac masses. For fixed  $(\mu_2, \dots, \mu_m)$ , let  $\mathbf{G}_1 \subseteq \mathcal{M}(\mathcal{X}_1)$  denote those measures that satisfy the constraints  $\mathbf{G}(\mu'_1, \mu_2, \dots, \mu_m) \leq 0$ . Consequently, since for  $k = 2, \dots, m$  and  $j = 1, \dots, n_k$ ,  $\mathbb{E}_{(\mu'_1, \mu_2, \dots, \mu_m)}[g_j^k]$  is constant in  $\mu'_1$ , it follows from [100, Theorem 2.1] that the extreme set  $\text{ex}(\mathbf{G}_1) \subseteq \mathcal{M}(\mathcal{X}_1)$  of the constraint set consists only of elements of  $\Delta_{n_1+n'}(\mathcal{X}_1)$ . In addition, von Weizsäcker and Winkler [99, Corollary 3] show that a Choquet theorem holds: let  $\mu'$  satisfy the constraints. Then

$$\mu'(B) = \int_{\text{ex}(\mathbf{G}_1)} \nu(B) \, dp(\nu),$$

for all Borel sets  $B \subseteq \mathcal{X}_1$ , where  $p$  is a probability measure on the extreme set  $\text{ex}(\mathbf{G}_1)$ .

According to Winkler, an extended-real-valued function  $K$  on  $\mathbf{G}_1$  is called *measure affine* if it satisfies the barycentric formula

$$K(\mu') = \int_{\text{ex}(\mathbf{G}_1)} K(\nu) \, dp(\nu).$$

When  $K$  is measure affine, [100, Theorem 3.2] asserts that

$$\sup_{\mu' \in \mathbf{G}_1} K(\mu') = \sup_{\nu \in \text{ex}(\mathbf{G}_1)} K(\nu),$$

and so we conclude that

$$\sup_{\mu' \in \mathbf{G}_1} K(\mu') = \sup_{\nu \in \text{ex}(\mathbf{G}_1)} K(\nu) \leq \sup_{\nu \in \Delta_{n_1+n'}(\mathcal{X}_1) \cap \mathbf{G}_1} K(\nu).$$

However, since

$$\sup_{\nu \in \Delta_{n_1+n'}(\mathcal{X}_1) \cap \mathbf{G}_1} K(\nu) \leq \sup_{\nu \in \mathbf{G}_1} K(\nu),$$

it follows that

$$\sup_{\mu' \in \mathbf{G}_1} K(\mu') = \sup_{\nu \in \Delta_{n_1+n'}(\mathcal{X}_1) \cap \mathbf{G}_1} K(\nu).$$



To apply this result, observe that [100, proposition 3.1] asserts that the evaluation function

$$\mu'_1 \mapsto \mathbb{E}_{\mu'_1} [\mathbb{E}_{(\mu_2, \dots, \mu_m)}[r]]$$

is measure affine. Therefore,

$$\sup_{\substack{\mu'_1 \in \mathcal{M}(\mathcal{X}_1) \\ \mathbf{G}(\mu'_1, \mu_2, \dots, \mu_m) \leq 0}} \mathbb{E}_{(\mu'_1, \mu_2, \dots, \mu_m)}[r] = \sup_{\substack{\mu'_1 \in \Delta_{n_1+n'}(\mathcal{X}_1) \\ \mathbf{G}(\mu'_1, \mu_2, \dots, \mu_m) \leq 0}} \mathbb{E}_{(\mu'_1, \mu_2, \dots, \mu_m)}[r]. \quad (10.1)$$

Now let  $\varepsilon > 0$  and let  $\mu_1^* \in \Delta_{n_1+n'}(\mathcal{X}_1)$  be  $\varepsilon$ -suboptimal for the right-hand side of (10.1): that is,  $\mathbf{G}(\mu_1^*, \mu_2, \dots, \mu_m) \leq 0$ , and

$$\mathbb{E}_{(\mu_1^*, \mu_2, \dots, \mu_m)}[r] \geq \sup_{\substack{\mu'_1 \in \Delta_{n_1+n'}(\mathcal{X}_1) \\ \mathbf{G}(\mu'_1, \mu_2, \dots, \mu_m) \leq 0}} \mathbb{E}_{(\mu'_1, \mu_2, \dots, \mu_m)}[r] - \varepsilon.$$

Hence, by (10.1),

$$\mathbb{E}_{(\mu_1^*, \mu_2, \dots, \mu_m)}[r] \geq \sup_{\substack{\mu'_1 \in \mathcal{M}(\mathcal{X}_1) \\ \mathbf{G}(\mu'_1, \mu_2, \dots, \mu_m) \leq 0}} \mathbb{E}_{(\mu'_1, \mu_2, \dots, \mu_m)}[r] - \varepsilon \geq \mathbb{E}_{(\mu_1, \mu_2, \dots, \mu_m)}[r] - \varepsilon.$$

Consequently, the first component of  $\mu$  by can be replaced some element of  $\Delta_{n_1+n'}(\mathcal{X}_1)$  to produce a feasible point  $\mu' \in \mathcal{M}^{\mathbf{G}}$  without decreasing  $\mathbb{E}[r]$  by more than  $\varepsilon$ . By repeating this argument, it follows that for every point  $\mu \in \mathcal{M}^{\mathbf{G}}$  there exists a  $\mu' \in \mathcal{M}_{\Delta}$  such that

$$\mathbb{E}_{\mu'}[r] \geq \mathbb{E}_{\mu}[r] - m\varepsilon.$$

Since  $\varepsilon$  was arbitrary the result follows.  $\square$

*Proof of Corollary 4.4.* Simply use the identity

$$\mathcal{U}(\mathcal{A}) = \sup_{(f, \mu) \in \mathcal{A}} \mathbb{E}_{\mu}[r_f] = \sup_{f \in \mathcal{G}} \sup_{\substack{\mu \in \mathcal{M}^m(\mathcal{X}) \\ \mathbf{G}(f, \mu) \leq 0}} \mathbb{E}_{\mu}[r_f]$$

and then apply Theorem 4.1 to the inner supremum.  $\square$

*Proof of Theorem 4.7.* Corollary 4.4 implies that  $\mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_{\Delta})$  where,

$$\mathcal{A}_{\Delta} := \left\{ (f, \mu) \in \mathcal{G} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i) \mid \mathbb{E}_{\mu}[g_i \circ f] \leq 0 \text{ for all } i = 1, \dots, n \right\}.$$

For each  $i = 1, \dots, m$ , the indexing of the Dirac masses pushes forward the measure  $\mu^i$  with weights  $\alpha_k^i, k = 0, \dots, n$  to a measure  $\alpha^i$  on  $\mathcal{N}$  with weights  $\alpha_k^i, k = 0, \dots, n$ . Let  $\alpha := \bigotimes_{i=1}^m \alpha^i$  denote the corresponding product measure on  $\mathcal{D} = \mathcal{N}^m$ . That is, we have a map

$$\mathbb{A}: \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i) \rightarrow \mathcal{M}^m(\mathcal{D})$$

and the product map

$$\mathbb{F} \times \mathbb{A}: \mathcal{G} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i) \rightarrow \mathcal{F}_{\mathcal{D}} \times \mathcal{M}^m(\mathcal{D}).$$

Since for any function  $g: \mathbb{R} \rightarrow \mathbb{R}$ , we have  $\mathbb{F}(g \circ f, \mu) = g \circ \mathbb{F}(f, \mu)$ , it follows that for any  $(f, \mu) \in \mathcal{F} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i)$  that

$$\mathbb{E}_{\mu}[g \circ f] = \mathbb{E}_{\alpha_{\mu}}[\mathbb{F}(g \circ f, \mu)] = \mathbb{E}_{\alpha_{\mu}}[g \circ \mathbb{F}(f, \mu)].$$

Consequently, with the function  $\mathcal{R}^{\mathcal{D}}: \mathcal{F}_{\mathcal{D}} \times \mathcal{M}^m(\mathcal{D}) \rightarrow \mathbb{R}$  defined by

$$\mathcal{R}^{\mathcal{D}}(h, \alpha) := \mathbb{E}_{\alpha}[r \circ h],$$

and for each  $j = 1, \dots, n$ , the functions  $\mathbf{G}_j^{\mathcal{D}}: \mathcal{F}_{\mathcal{D}} \times \mathcal{M}^m(\mathcal{D}) \rightarrow \mathbb{R}$  defined by

$$\mathbf{G}_j^{\mathcal{N}}(h, \alpha) := E_{\alpha}[g_i \circ h],$$

we have that, for all  $(f, \mu) \in \mathcal{F} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i)$ ,

$$\mathcal{R}(f, \mu) = \mathcal{R}^{\mathcal{D}}(\mathbb{F}(f, \mu), \alpha_{\mu}), \quad (10.2)$$

and, for all  $j = 1, \dots, n$  and all  $(f, \mu) \in \mathcal{F}_{\mathcal{D}} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i)$ ,

$$\mathbf{G}_j(f, \mu) = \mathbf{G}_j^{\mathcal{D}}(\mathbb{F}(f, \mu), \alpha_{\mu}), \quad (10.3)$$

where  $\alpha_{\mu} := \mathbb{A}(\mu)$ .

That is,

$$\mathcal{R} = \mathcal{R}^{\mathcal{D}} \circ (\mathbb{F} \times \mathbb{A}),$$

$$\mathbf{G}_j = \mathbf{G}_j^{\mathcal{D}} \circ (\mathbb{F} \times \mathbb{A}) \text{ for each } j = 1, \dots, n.$$

Consequently, any  $(f, \mu) \in \mathcal{A}_{\Delta}$  is mapped by  $\mathbb{F} \times \mathbb{A}$  to a point in  $\mathcal{F} \times \mathcal{M}^m(\mathcal{D})$  that preserves the criterion value and the constraint, and by the assumption must lie in  $\mathcal{G}_{\mathcal{D}} \times \mathcal{M}^m(\mathcal{D})$ . This establishes  $\mathcal{U}(\mathcal{A}_{\Delta}) \leq \mathcal{U}(\mathcal{A}_{\mathcal{D}})$ .

To obtain equality, consider  $(h, \alpha) \in \mathcal{A}_{\mathcal{D}}$ . By assumption, there exists an  $(f, \mu) \in \mathcal{G} \times \bigotimes_{i=1}^m \Delta_n(\mathcal{X}_i)$  such that  $\mathbb{F}(f, \mu) = h$ . If we adjust the weights on  $\mu$  so that  $\mathbb{A}(\mu) = \alpha$ , we still maintain  $\mathbb{F}(f, \mu) = h$ . By (10.2) and (10.3), this point has the same criterion value and satisfies the integral constraints of  $\mathcal{A}_{\Delta}$ . The proof is finished.  $\square$

*Proof of Proposition 4.8.* Let  $I := \mathbb{1}_{[a, \infty)}$  be the indicator function and consider  $r_f := I \circ f$  so that  $\mu[f \geq a] = \mathbb{E}_{\mu}[I \circ f]$ . Since  $I \circ f$  is integrable for all  $\mu \in \mathcal{M}^m(\mathcal{X})$  and we have one constraint  $\mathbb{E}_{\mu}[f] \leq 0$ , the result follows from Theorem 4.7, provided that we have

$$\mathbb{F} \left( \mathcal{G} \times \bigotimes_{i=1}^m \Delta_1(\mathcal{X}_i) \right) = \mathcal{G}_{\mathcal{D}}.$$

To establish this, consider  $f \in \mathcal{G}$  and observe that for all  $\mu \in \bigotimes_{i=1}^m \Delta_1(\mathcal{X}_i)$  it holds that  $\mathbb{F}(f, \mu) \in \mathcal{G}_{\mathcal{D}}$ . Therefore, we conclude that  $\mathbb{F}(\mathcal{G} \times \bigotimes_{i=1}^m \Delta_1(\mathcal{X}_i)) \subseteq \mathcal{G}_{\mathcal{D}}$ . On the other hand, for any  $h \in \mathcal{G}_{\mathcal{D}}$ , we can choose a measurable product partition of  $\mathcal{X}$  dividing each  $\mathcal{X}_i$  into 2 cells. We pull back the function  $h$  to a function  $f \in \mathcal{F}$  that has the correct constant values in the partition cells, and place the Dirac masses into the correct cells. Set the weights to any nonzero values. It is easy to see that  $f \in \mathcal{G}$ . Moreover, we have a measure  $\mu$  which satisfies  $\mathbb{F}(f, \mu) = h$ . Therefore, we conclude that  $\mathbb{F}(\mathcal{G} \times \bigotimes_{i=1}^m \Delta_1(\mathcal{X}_i)) \supseteq \mathcal{G}_{\mathcal{D}}$ . This completes the proof.  $\square$

*Proof of Theorem 4.9.* First, observe that  $\mathcal{G}_{\mathcal{D}}$  is a sub-lattice of  $\mathcal{F}_{\mathcal{D}}$  in the usual lattice structure on functions. That is, if  $h_1, h_2 \in \mathcal{G}_{\mathcal{D}}$ , then it follows that both  $\min(h_1, h_2) \in \mathcal{G}_{\mathcal{D}}$  and  $\max(h_1, h_2) \in \mathcal{G}_{\mathcal{D}}$ . Therefore, for any admissible  $(h, \alpha) \in \mathcal{A}_{\mathcal{D}}$ , it follows that clipping  $h$  at  $a$  to  $\min(h, a)$  produces an admissible  $(\min(h, a), \alpha)$  and does not change the criterion value  $\alpha[h \geq a]$ . Consequently, we can reduce to functions with maximum value  $a$ . Moreover, since each function  $h^s$  is in the sub-lattice  $\mathcal{G}_{\mathcal{D}}$ , it follows that  $h^C \in \mathcal{G}_{\mathcal{D}}, C \in \mathcal{C}$ . For  $C \in \mathcal{C}$ , define the sub-lattice

$$C_{\mathcal{D}} := \{h \in \mathcal{F}_{\mathcal{D}} \mid \{s \mid h(s) = a\} = C\}$$

of functions with value  $a$  precisely on the set  $C$ . Now, consider a function  $h \in \mathcal{G}_{\mathcal{D}}$  such that  $h \leq a$  and let  $C$  be the set where  $h = a$ . It follows that  $h^C \leq h$ ,  $h^C \in \mathcal{G}_{\mathcal{D}}$ , and  $h^C \in C_{\mathcal{D}}$ . Since  $h^C \leq h$  implies that  $\mathbb{E}_{\alpha}[h^C] \leq \mathbb{E}_{\alpha}[h]$  for all  $\alpha$ , it follows that replacing  $(h, \alpha)$  by  $(h^C, \alpha)$  keeps it admissible, and  $\alpha[h^C \geq a] = \alpha[h \geq a]$ . The proof is finished.  $\square$

## 10.2 Proofs for Section 5

The proofs given in this subsection are direct applications of Theorem 4.9. In particular, they are based on an analytical calculation of (4.19). Because Proposition 5.7 is fundamental to all the other results of the section, its proof will be given first.

*Proof of Proposition 5.7.* When non-ambiguous, we will use the notation  $\mathbb{E}[h^{C_0}]$  for  $\mathbb{E}_{\alpha}[h^{C_0}]$  and  $\mathbb{P}[h^{C_0} \geq a]$  for  $\alpha[h^{C_0} \geq a]$ . First, observe that, if  $\sum_{j=1}^m D_j \leq a$ , then  $\min(h^{C_0}) \geq 0$ , and the constraint  $\mathbb{E}[h^{C_0}] \leq 0$  imply  $\mathbb{P}[h^{C_0} = 0] = 1$ . This proves the first equation of (5.9). Now, assume  $a < \sum_{j=1}^m D_j$  and observe that

$$h^{C_0}(s) = a - \sum_{j=1}^m (1 - s_j) D_j.$$

It follows that

$$\mathbb{E}_{\alpha}[h^{C_0}] = a - \sum_{j=1}^m (1 - \alpha_j) D_j. \quad (10.4)$$

If  $D_m = 0$ , then the optimum is achieved on boundary of  $[0, 1]^m$  (i.e. by taking  $\alpha_m = 1$  since  $C_0 = \{(1, \dots, 1)\}$  and since  $h^{C_0}$  does not depend on  $s_m$ ) and the optimization

reduces to an  $(m - 1)$ -dimensional problem. For that reason, we will assume in all of the proofs of the results given in this section that all the  $D_i$ s are strictly positive. The statements of those results remain valid even if one or more of the  $D_i$ s are equal to zero.

The condition  $D_m > 0$  implies that  $\min(D_1, \dots, D_m) > 0$  and that

$$\alpha[h^{C_0} \geq a] = \prod_{j=1}^m \alpha_j. \quad (10.5)$$

If the optimum in  $\alpha$  is achieved in the interior of the hypercube  $[0, 1]^m$ , then at that optimum the gradients of (10.4) and (10.5) are collinear. Hence, in that case, there exists  $\lambda \in \mathbb{R}$  such that for all  $i \in \{1, \dots, m\}$ ,

$$\frac{\prod_{j=1}^m \alpha_j}{\alpha_i} = \lambda D_i. \quad (10.6)$$

Since  $\alpha[h^{C_0} \geq a]$  is increasing in each  $\alpha_j$ , the optimum is achieved at  $\mathbb{E}_\alpha[h^{C_0}] = 0$ . Substitution of (10.6) into the equation  $\mathbb{E}_\alpha[h^{C_0}] = 0$  yields that

$$\lambda = \frac{m \prod_{j=1}^m \alpha_j}{\sum_{j=1}^m D_j - a}$$

and, hence,

$$\alpha_i = \frac{\sum_{j=1}^m D_j - a}{m D_i}. \quad (10.7)$$

For all  $i \in \{1, \dots, m\}$ , the condition  $0 < \alpha_i < 1$  is equivalent to  $a < \sum_{j=1}^m D_j$  and

$$\sum_{j=1}^m D_j < a + m D_i. \quad (10.8)$$

It follows that for  $\sum_{j=1}^m D_j - m D_m < a < \sum_{j=1}^m D_j$ , the  $\alpha$  defined by (10.7) lies in the interior of  $[0, 1]^m$  and satisfies

$$\alpha[h^{C_0} \geq a] = \frac{\left(\sum_{j=1}^m D_j - a\right)^m}{m^m \prod_{j=1}^m D_j}.$$

If  $a \leq \sum_{j=1}^m D_j - m D_m$ , then the optimum is achieved on boundary of  $[0, 1]^m$  (i.e. by taking  $\alpha_m = 1$ , since  $C_0 = \{(1, \dots, 1)\}$ ), and the optimization reduces to an  $(m - 1)$ -dimensional problem.

To complete the proof, we use an induction. Observe in particular that, for  $k \leq m - 1$ ,

$$\frac{(\sum_{j=1}^k D_j - a)^k}{k^k \prod_{j=1}^k D_j} = \frac{(\sum_{j=1}^{k+1} D_j - a)^{k+1}}{(k+1)^{k+1} \prod_{j=1}^{k+1} D_j}$$

for  $a = \sum_{j=1}^{k+1} D_j - (k+1)D_{k+1}$ , and that

$$\frac{(\sum_{j=1}^k D_j - a)^k}{k^k \prod_{j=1}^k D_j} \leq \frac{(\sum_{j=1}^{k+1} D_j - a)^{k+1}}{(k+1)^{k+1} \prod_{j=1}^{k+1} D_j} \quad (10.9)$$

is equivalent to  $a \geq \sum_{j=1}^{k+1} D_j - (k+1)D_{k+1}$ . Indeed, writing  $a = \sum_{j=1}^{k+1} D_j - (k+1)D_{k+1} + b$ , equation (10.9) is equivalent to

$$\left(1 - \frac{b}{kD_{k+1}}\right)^k \leq \left(1 - \frac{b}{(k+1)D_{k+1}}\right)^{k+1}.$$

The function  $f$  given by  $f(x) := \left(1 - \frac{y}{x}\right)^x$  is increasing in  $x$  (for  $0 < y < x$ ): simply examine the derivative of  $\log f$ , and use the elementary inequality

$$\log(1-z) + \frac{z}{1-z} \geq 0 \text{ for } 0 < z < 1.$$

We will now give the outline of the induction. It is trivial to obtain that equation (5.9) is true for  $m = 1$ . Assume that it is true for  $m = q-1$  and consider the case  $m = q$ . Equation (10.7) isolates the only potential optimizer  $\alpha^q$ , which is not on the boundary of  $[0, 1]^q$  (not  $(q-1)$ -dimensional).

One obtains that equation (5.9) holds for  $m = q$  by comparing the value of  $\alpha[h^{C_0} \geq a]$  at locations  $\alpha$  isolated by equations (10.7) and (10.8) with those isolated at step  $q-1$ . This comparison is performed via equation (10.9).

More precisely, if  $\alpha^q$  (the candidate for the optimizer in  $\alpha$  isolated by the previous paragraph) is not an optimum, then the optimum must lie in the boundary of  $[0, 1]^q$ . Hence, the optimum must be achieved by taking  $\alpha_i = 1$  for some  $i \in \{1, \dots, q\}$ . Observing that  $\mathcal{U}(\mathcal{A}_{C_0})$  is increasing in each  $D_i$ , and since  $D_q = \min_{i \in \{1, \dots, q\}} D_i$ , that optimum can be achieved by taking  $i = q$ , which leads to computing  $\mathcal{U}(\mathcal{A}_{C_0})$  with  $(D_1, \dots, D_{q-1})$ , where we can use the  $(q-1)$ -step of the induction. Using equation (10.9) for  $k = q-1$ , we obtain that  $\alpha^q$  is an optimum for  $a \geq \sum_{j=1}^q D_j - qD_q$ , and that, for  $a \leq \sum_{j=1}^q D_j - qD_q$ , the optimum is achieved by calculating  $\mathcal{U}(\mathcal{A}_{C_0})$  with  $q-1$  variables and  $(D_1, \dots, D_{q-1})$ . This finishes the proof by using the induction assumption (see formula (5.9)).  $\square$

The following two lemmas illustrate simplifications that can be made using the symmetries of the hypercube:

**Lemma 10.1.** *Let  $C_0 \in \mathcal{C}$ . If  $C_0$  is symmetric with respect to the hyperplane containing the center of the hypercube and normal to the direction  $i$ , then the optimum of (5.8) can be achieved by taking  $\alpha_i = 1$ .*

*Proof.* The proof follows by observing that if  $C_0$  is symmetric with respect to the hyperplane containing the center of the hypercube and normal to the direction  $i$ , then  $h^{C_0}(s)$  does not depend on the variable  $s_i$ .  $\square$

The following lemma is trivial:

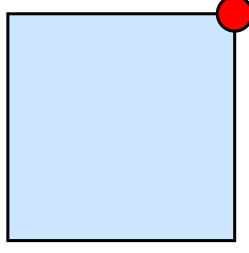


Figure 10.1: For  $m = 2$ , the optimum associated with  $\mathcal{U}(\mathcal{A}_{\mathcal{C}})$  can be achieved with  $C = \{(1, 1)\}$ . For that specific value of  $C$ , the linearity of  $h^C(s) = a - D_1(1 - s_1) - D_2(1 - s_2)$  implies  $\mathcal{U}(\mathcal{A}_{\text{Hfd}}) = \mathcal{U}(\mathcal{A}_{\text{McD}})$ .

**Lemma 10.2.** *Let  $(\alpha, C)$  be an optimizer of (4.19). Then, the images of  $(\alpha, C)$  by reflections with respect to hyperplanes containing the center of the hypercube and normal to its faces are also optimizers of (4.19).*

The proofs of the remaining theorems now follow in the order that the results were stated in the main part of the paper.

*Proof of Theorem 5.1.* The calculation of  $\mathcal{U}(\mathcal{A}_{\mathcal{C}})$  for  $m = 1$  is trivial and also follows from Proposition 5.7.  $\square$

*Proof of Theorem 5.2.* Write  $C_1 = \{(1, 1)\}$  (see Figure 10.1). Theorem 5.2 is a consequence of the following inequality:

$$\max_{C_0 \in \mathcal{C}} \mathcal{U}(\mathcal{A}_{C_0}) \leq \mathcal{U}(\mathcal{A}_{C_1}) \quad (10.10)$$

Assuming equation (10.10) to be true, equation (5.3) is obtained by calculating  $\mathcal{U}(\mathcal{A}_{C_1})$  from Proposition 5.7 with  $m = 2$ . Let us now prove equation (10.10). Let  $C_0 \in \mathcal{C}$ ; we need to prove that

$$\mathcal{U}(\mathcal{A}_{C_0}) \leq \mathcal{U}(\mathcal{A}_{C_1}). \quad (10.11)$$

By symmetry (using Lemma 10.2), it is no loss of generality to assume that  $(1, 1) \in C_0$ . By Lemma 10.1 the optima for  $C_0 = \{(1, 1), (1, 0)\}$  and  $C_0 = \{(1, 1), (0, 1)\}$  can be achieved with  $C_1$  by taking  $\alpha_1 = 1$  or  $\alpha_2 = 1$ .

The case  $C_0 = \{(1, 1); (1, 0); (0, 1); (0, 0)\}$  is infeasible.

For  $C_0 = \{(1, 1), (1, 0), (0, 1)\}$ , we have  $\mathbb{P}[h^{C_0} = a] = \beta$  and  $\mathbb{E}[h^{C_0}] = a - (1 - \beta) \min(D_1, D_2)$  with  $\beta = 1 - (1 - \alpha_1)(1 - \alpha_2)$  (recall that  $h^{C_0}$  is defined by equation (4.17)). Hence, at the optimum (in  $\alpha$ ),

$$\mathbb{P}[h^{C_0} = a] = \begin{cases} 1 - a / \min(D_1, D_2), & \text{if } a < \min(D_1, D_2), \\ 0, & \text{if } a \geq \min(D_1, D_2). \end{cases} \quad (10.12)$$

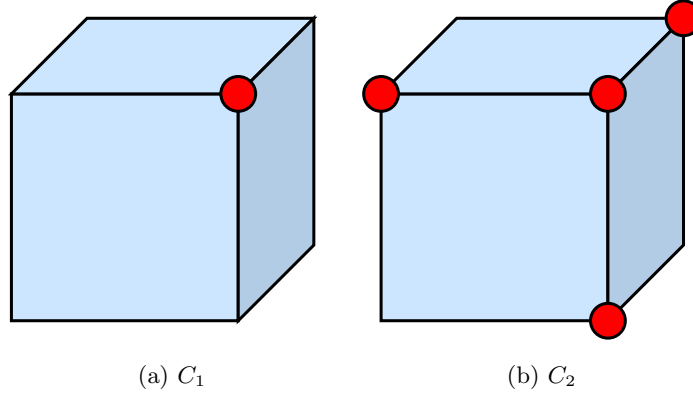


Figure 10.2: For  $m = 3$ , the optimum associated with  $\mathcal{U}(\mathcal{A}_C)$  can be achieved with  $C_1 = \{(1, 1, 1)\}$  (leading to  $\mathcal{F}_1$ ) or  $C_2 = \{(1, 1, 1), (0, 1, 1), (1, 0, 1), (1, 1, 0)\}$  (leading to  $\mathcal{F}_2$ ). The linearity of  $h^{C_1}(s) = a - D_1(1 - s_1) - D_2(1 - s_2) - D_3(1 - s_3)$  implies that  $\mathcal{U}(\mathcal{A}_{\text{Hfd}}) = \mathcal{U}(\mathcal{A}_{\text{McD}})$  when  $\mathcal{F}_1 \geq \mathcal{F}_2$ . Similarly, the nonlinearity of  $h^{C_2}$  leads to  $\mathcal{U}(\mathcal{A}_{\text{Hfd}}) < \mathcal{U}(\mathcal{A}_{\text{McD}})$  when  $\mathcal{F}_1 < \mathcal{F}_2$ .

Equation (10.11) then holds by observing that one always has both

$$1 - \frac{a}{\min(D_1, D_2)} \leq 1 - \frac{a}{\max(D_1, D_2)}$$

and

$$1 - \frac{a}{\min(D_1, D_2)} \leq \frac{(D_1 + D_2 - a)^2}{4D_1D_2}.$$

The last inequality is equivalent to  $(D_1 - D_2 + a)^2 \geq 0$ , which is always true. The case  $C_0 = \{(1, 1), (0, 0)\}$  is bounded by the previous one since  $\mathbb{P}[h^{C_0} = a] = \beta$  and  $\mathbb{E}[h^{C_0}] = a\beta - (1 - \beta)\min(D_1, D_2)$  with  $\beta = \alpha_1\alpha_2 + (1 - \alpha_1)(1 - \alpha_2)$ . This finishes the proof.  $\square$

*Proof of Theorem 5.4.* Recall that

$$\mathcal{U}(\mathcal{A}_{\text{McD}}) = \max_{C_0 \in \mathcal{C}} \mathcal{U}(\mathcal{A}_{C_0}).$$

It follows from Proposition 5.7 that  $\mathcal{F}_1$  corresponds to  $\mathcal{U}(\mathcal{A}_{C_1})$  with  $C_1 = \{(1, 1, 1)\}$ . Write  $C_2 = \{(1, 1, 1), (0, 1, 1), (1, 0, 1), (1, 1, 0)\}$  (see Figure 10.2). Let us now calculate  $\mathcal{U}(\mathcal{A}_{C_2})$  ( $\mathcal{F}_2$  corresponds to  $\mathcal{U}(\mathcal{A}_{C_2})$ , which is the optimum, and it is achieved in the interior of  $[0, 1]^3$ ). We have  $\mathbb{P}[h^{C_2} = a] = \alpha_2\alpha_3 + \alpha_1\alpha_3 + \alpha_1\alpha_2 - 2\alpha_1\alpha_2\alpha_3$ , and

$$\mathbb{E}[h^{C_2}] = a - D_2(1 - \alpha_1)(1 - \alpha_2) - D_3((1 - \alpha_2)(1 - \alpha_3) + (1 - \alpha_1)\alpha_2(1 - \alpha_3)).$$

An internal optimal point  $\alpha$  must satisfy, for some  $\lambda \in \mathbb{R}$ ,

$$\alpha_2 + \alpha_3 - 2\alpha_2\alpha_3 = \lambda(D_2(1 - \alpha_2) + D_3\alpha_2(1 - \alpha_3)), \quad (10.13a)$$

$$\alpha_1 + \alpha_3 - 2\alpha_1\alpha_3 = \lambda(D_2(1 - \alpha_1) + D_3\alpha_1(1 - \alpha_3)), \quad (10.13b)$$

$$\alpha_1 + \alpha_2 - 2\alpha_1\alpha_2 = \lambda(D_3(1 - \alpha_1\alpha_2)). \quad (10.13c)$$

If we multiply the first equation by  $\alpha_1$  and subtract the second equation multiplied by  $\alpha_2$ , then we obtain that

$$(\alpha_1 - \alpha_2)\alpha_3 = \lambda D_2(\alpha_1 - \alpha_2),$$

which implies that either  $\alpha_1 = \alpha_2$  or  $\alpha_3 = \lambda D_2$ .

Suppose that  $\alpha_1 \neq \alpha_2$ , so that  $\alpha_3 = \lambda D_2$ . Subtraction of the second equation in (10.13) from the first yields

$$(\alpha_2 - \alpha_1)(1 - 2\alpha_3) = \lambda(-D_2(\alpha_2 - \alpha_1) + D_3(\alpha_2 - \alpha_1)(1 - \alpha_3)),$$

which implies that either  $\alpha_1 = \alpha_2$  or

$$1 - 2\alpha_3 = \lambda(-D_2 + D_3(1 - \alpha_3)).$$

Since  $\alpha_3 = \lambda D_2$ , this becomes

$$1 - \alpha_3 = \frac{\alpha_3 D_3}{D_2}(1 - \alpha_3),$$

which implies, since  $\alpha_3 \neq 1$ , that  $\alpha_3 = \frac{D_2}{D_3}$ . Therefore,  $\lambda = \frac{1}{D_3}$ . Therefore, the third equation in (10.13) becomes

$$\alpha_1 + \alpha_2 - 2\alpha_1\alpha_2 = \lambda(D_3(1 - \alpha_1\alpha_2)) = 1 - \alpha_1\alpha_2,$$

which amounts to

$$\alpha_1 + \alpha_2 - \alpha_1\alpha_2 = 1,$$

which in turn amounts to  $\alpha_1(1 - \alpha_2) = 1 - \alpha_2$ . Since  $\alpha_2 \neq 1$ , we conclude that  $\alpha_1 = 1$ , contradicting the supposition that  $\alpha$  is an interior point. Therefore,  $\alpha_1 = \alpha_2$  and equations (10.13), with  $\alpha := \alpha_1 = \alpha_2$ , become

$$\alpha + \alpha_3 - 2\alpha\alpha_3 = \lambda(D_2(1 - \alpha) + D_3\alpha(1 - \alpha_3)) \quad (10.14a)$$

$$2\alpha - 2\alpha^2 = \lambda(D_3(1 - \alpha^2)). \quad (10.14b)$$

Hence,

$$\mathbb{P}[h^{C_2} = a] = 2\alpha\alpha_3 + \alpha^2 - 2\alpha^2\alpha_3 \quad (10.15)$$

and

$$\mathbb{E}[h^{C_2}] = a - D_2(1 - \alpha)^2 - D_3((1 - \alpha^2)(1 - \alpha_3)).$$

The hypothesis that the optimum is not achieved on the boundary requires that

$$D_3, 0 < \alpha < 1, D_2 + D_3 > a \text{ and } E[h^{C_2}] = 0.$$



The condition  $\mathbb{E}[h^{C_2}] = 0$  is required because equation (10.15) is strictly increasing along the direction  $\alpha = \alpha_3$ .

Suppose that those conditions are satisfied. The condition  $\mathbb{E}[h^{C_2}] = 0$  implies that

$$1 - \alpha_3 = \frac{a}{D_3(1 - \alpha^2)} - \frac{D_2(1 - \alpha)}{D_3(1 + \alpha)},$$

which in turns implies that

$$\alpha_3 = 1 - \frac{a}{D_3(1 - \alpha^2)} + \frac{D_2(1 - \alpha)}{D_3(1 + \alpha)}. \quad (10.16)$$

Substitution of (10.16) into (10.15) yields that  $\mathbb{P}[h^{C_2} = a] = \Psi(\alpha)$ , with

$$\Psi(\alpha) = \alpha^2 + 2(\alpha - \alpha^2) \left( 1 - \frac{a}{D_3(1 - \alpha^2)} + \frac{D_2(1 - \alpha)}{D_3(1 + \alpha)} \right).$$

Hence,

$$\Psi(\alpha) = 2\alpha - \alpha^2 - 2\frac{a}{D_3} \frac{\alpha}{1 + \alpha} + 2\frac{D_2}{D_3} \alpha \frac{(1 - \alpha)^2}{1 + \alpha}.$$

$\Psi(\alpha)$  can be simplified using polynomial division. In particular, using

$$\alpha \frac{(1 - \alpha)^2}{1 + \alpha} = (1 - \alpha)^2 - \frac{(1 - \alpha)^2}{1 + \alpha},$$

$$\alpha \frac{(1 - \alpha)^2}{1 + \alpha} = \alpha^2 + 1 - 2\alpha - (1 + \alpha) + 4 - \frac{4}{1 + \alpha},$$

where the last step is obtained from

$$(1 - \alpha)^2 = (\alpha + 1 - 2)^2 = (\alpha + 1)^2 - 4(1 + \alpha) + 4,$$

we obtain that

$$\Psi(\alpha) = 2\alpha - \alpha^2 - 2\frac{a}{D_3} \frac{\alpha}{1 + \alpha} + 2\frac{D_2}{D_3} \left( 4 + \alpha^2 - 3\alpha - \frac{4}{1 + \alpha} \right).$$

Therefore,

$$\Psi(\alpha) = \alpha^2 \left( 2\frac{D_2}{D_3} - 1 \right) - \frac{2a}{D_3} \frac{\alpha}{1 + \alpha} + 2\alpha \left( 1 - 3\frac{D_2}{D_3} \right) + 8\frac{D_2}{D_3} \frac{\alpha}{1 + \alpha}$$

and

$$\Psi(\alpha) = \alpha^2 \left( 2\frac{D_2}{D_3} - 1 \right) - 2\alpha \left( 3\frac{D_2}{D_3} - 1 \right) + \frac{\alpha}{1 + \alpha} \left( 8\frac{D_2}{D_3} - 2\frac{a}{D_3} \right). \quad (10.17)$$

Equation (10.17) implies that

$$D_3\Psi'(\alpha) = 2\alpha(2D_2 - D_3) + 2(D_3 - 3D_2) - \frac{1}{(1 + \alpha)^2}(2a - 8D_2).$$

The equation  $\Psi'(\alpha) = 0$  is equivalent to equation (5.6). An interior optimum requires the existence of an  $\alpha \in (0, 1)$  such that  $\Psi'(\alpha) = 0$  and  $\alpha_3 \in (0, 1)$ , which leads to the definition of  $\mathcal{F}_2$ . This establishes the theorem for the  $\mathcal{F}_2$  case.

Next, using symmetries of the hypercube and through direct computation (as in the  $m = 2$  case), we obtain that

$$C_0 \neq C_2 \implies \mathcal{U}(\mathcal{A}_{C_0}) \leq \mathcal{U}(\mathcal{A}_{C_1}). \quad (10.18)$$

For the sake of concision, we will give the detailed proof of (10.18) only for

$$C_3 = \{(1, 1, 1), (0, 1, 1), (1, 0, 1)\}.$$

This proof will give an illustration of generic reduction properties used in other cases. To address all the symmetric transformations of  $C_3$ , we will give the proof without assuming that  $D_1, D_2$  and  $D_3$  are ordered. Let us now consider the  $C_3$  scenario. If the optimum in  $\alpha$  is achieved on the boundary of  $[0, 1]^3$ , then equation (10.10) implies  $\mathcal{U}(\mathcal{A}_{C_3}) \leq \mathcal{U}(\mathcal{A}_{C_1})$ . Let us assume that the optimum is not achieved on the boundary of  $[0, 1]^3$ . Observe that

$$h^{C_3}(s_1, s_2, 0) = h^{C_3}(s_1, s_2, 1) - D_3. \quad (10.19)$$

Combining (10.19) with

$$\mathbb{E}[h^{C_3}] = \alpha_3 \mathbb{E}[h^{C_3}(s_1, s_2, 1)] + (1 - \alpha_3) \mathbb{E}[h^{C_3}(s_1, s_2, 0)]$$

implies that

$$\mathbb{E}[h^{C_3}] = \mathbb{E}[h^{C_3}(s_1, s_2, 1)] - (1 - \alpha_3)D_3.$$

Furthermore,

$$\mathbb{P}[h^{C_3} = a] = \alpha_3 \mathbb{P}[h^{C_3}(s_1, s_2, 1) = a]. \quad (10.20)$$

Maximizing (10.20) with respect to  $\alpha_3$  under the constraint  $\mathbb{E}[h^{C_3}] \leq 0$  leads to  $\mathbb{E}[h^{C_3}] = 0$  (because  $\mathbb{P}[h^{C_3} = a]$  and  $\mathbb{E}[h^{C_3}]$  are linear in  $\alpha_3$ ) and

$$\alpha_3 = 1 - \frac{\mathbb{E}[h^{C_3}(s_1, s_2, 1)]}{D_3}. \quad (10.21)$$

Observe that the condition  $\alpha_3 < 1$  requires  $\mathbb{E}[h^{C_3}(s_1, s_2, 1)] > 0$ . If  $\mathbb{E}[h^{C_3}(s_1, s_2, 1)] \leq 0$  then  $\alpha_3 = 1$ , and the optimum is achieved on the boundary of  $[0, 1]^3$ .

The maximization of  $\mathbb{P}[h^{C_3}(s_1, s_2, 1) = a]$  under the constraint  $\mathbb{E}[h^{C_3}(s_1, s_2, 1)] \leq E$  (where  $E$  is a slack optimization variable) leads to (using the  $m = 2$  result)

$$\mathbb{P}[h^{C_3}(s_1, s_2, 1) = a] = 1 - \frac{(a - E)}{\min(D_1, D_2)}$$

if  $a - E \leq \min(D_1, D_2)$ , and  $\mathbb{P}[h^{C_3}(s_1, s_2, 1) = a] = 0$  otherwise. It follows from (10.21) and (10.20) that if the optimum is achieved at an interior point, then the optimal value of  $\mathbb{P}[h^{C_3} = a]$  is achieved by taking the maximum of

$$\mathbb{P}[h^{C_3} = a] = \left(1 - \frac{E}{D_3}\right) \left(1 - \frac{a - E}{\min(D_1, D_2)}\right)$$

with respect to  $E$  with the constraints  $0 \leq E \leq D_3$  and  $a - \min(D_1, D_2) \leq E$ . If the optimum is not achieved on the boundary of  $[0, 1]^3$  then one must have

$$E = \frac{D_3 + \min(D_1, D_2) - a}{2},$$

which leads to

$$\mathbb{P}[h^{C_3} = a] = \frac{(D_3 + \min(D_1, D_2) - a)^2}{4D_3 \min(D_1, D_2)}. \quad (10.22)$$

Comparison of (10.22) and (5.3) implies that  $\mathcal{U}(\mathcal{A}_{C_3}) \leq \mathcal{U}(\mathcal{A}_{C_1})$ , by Proposition 5.7.  $\square$

*Proof of Proposition 5.9.* The idea of the proof is to show that  $h^C$  can be chosen so that  $C$  contains only one vertex of the hypercube, in which case we have the explicit formula obtained in Proposition 5.7.

First, observe that if  $a > \sum_{j=1}^{m-1} D_j$ , then it is not possible to satisfy the constraint  $\mathbb{E}_\alpha[h^C] \leq 0$  whenever  $C$  contains two or more vertices of the hypercube. Next, if  $C$  contains two vertices  $s^1, s^2$  of the hypercube, and the Hamming distance between those points is 1, then  $C$  is symmetric with respect to a hyperplane containing the center of the hypercube and normal to one of its faces, and the problem reduces to dimension  $m-1$ . It follows from Lemma 10.1 that the optimum of (5.8) can be achieved by a  $C$  that has only one element. If  $C$  contains two vertices of the hypercube, and the Hamming distance between those points is greater than or equal to 2, then the constraint  $\mathbb{E}_\alpha[h^C] \leq 0$  is infeasible if  $a > \sum_{j=1}^{m-2} D_j + D_m$  (because  $h^C > 0$  in that case). Therefore, we conclude using Proposition 5.7.  $\square$

*Proof of Theorem 5.11.* First, we observe that we always have

$$\mathcal{U}(\mathcal{A}_{\text{Hfd}}) \leq \mathcal{U}(\mathcal{A}_{\text{McD}}). \quad (10.23)$$

We observe from equation (10.10) that the maximizer ( $h^C$ ) of  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  is linear (see Figure 10.1), and hence is also an admissible function under  $\mathcal{U}(\mathcal{A}_{\text{Hfd}})$ . This finishes the proof.  $\square$

*Proof of Theorem 5.13.* Just as for  $m = 2$ , equation (10.23) is always satisfied. Next, observing that  $\mathcal{F}_1$ , in Theorem 5.4, is associated with a linear maximizer  $h^C$  (see Figure 10.2), we deduce that

$$\mathcal{F}_1 \leq \mathcal{U}(\mathcal{A}_{\text{Hfd}}) \leq \max(\mathcal{F}_1, \mathcal{F}_2).$$

This finishes the proof for equation (5.12). Let us now prove equation (5.13). Assuming that  $\mathcal{U}(\mathcal{A}_{\text{Hfd}}) = \mathcal{U}(\mathcal{A}_{\text{McD}})$ , it follows that  $\mathcal{U}(\mathcal{A}_{\text{McD}})$  can be achieved by a linear function  $h_0$ . Since at the optimum we must have  $\mathbb{E}[h_0] = 0$ , and since  $\min(h_0, a)$  is also a maximizer of  $\mathcal{U}(\mathcal{A}_{\text{McD}})$ , it follows that  $\min(h_0, a) = h_0$ . Using the linearity of  $h_0$ , and the lattice structure of the set of functions in  $\mathcal{U}(\mathcal{A}_{\text{McD}})$ , we deduce that  $h_0 = h^C$ , where  $C$  contains only one vertex of the cube. It follows that  $\mathcal{F}_1 \geq \mathcal{F}_2$ . This finishes the proof.  $\square$

## References

- [1] M. Adams, A. Lashgari, B. Li, M. McKerns, J. Mihaly, M. Ortiz, H. Owhadi, A.J. Rosakis, M. Stalzer, and T.J. Sullivan. Rigorous model-based uncertainty quantification with application to terminal ballistics, part ii: Systems with uncontrollable inputs and large scatter. *Journal of the Mechanics and Physics of Solids*, 60(5):1002–1019, 2012.
- [2] C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis: A Hitchhiker’s Guide*. Springer, Berlin, third edition, 2006.
- [3] T. W. Anderson. The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.*, 6(2):170–176, 1955.
- [4] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Rev.*, 52(2):317–355, 2010.
- [5] R. E. Barlow and F. Proschan. *Mathematical Theory of Reliability*. SIAM, Philadelphia, 1996.
- [6] J. L. Beck. Bayesian system identification based on probability logic. *Struct. Control Health Monit.*, 17:825–847, 2010.
- [7] J. L. Beck and L. S. Katafygiotis. Updating models and their uncertainties: Bayesian statistical framework. *Journal of Engineering Mechanics*, 124(4):455–461, 1998.
- [8] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures*, volume 5 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1978.
- [9] V. Bentkus. A remark on the inequalities of Bernstein, Prokhorov, Bennett, Hoeffding, and Talagrand. *Liet. Mat. Rink.*, 42(3):332–342, 2002.
- [10] V. Bentkus. On Hoeffding’s inequalities. *Ann. Probab.*, 32(2):1650–1673, 2004.
- [11] V. Bentkus, G. D. C. Geuze, and M. C. A. van Zuijlen. Optimal Hoeffding-like inequalities under a symmetry assumption. *Statistics*, 40(2):159–164, 2006.
- [12] S. N. Bernstein. *Collected Works*. Izdat. “Nauka”, Moscow, 1964.
- [13] D. Bertsimas and I. Popescu. Optimal inequalities in probability theory: a convex optimization approach. *SIAM J. Optim.*, 15(3):780–804 (electronic), 2005.
- [14] M. Bieri and C. Schwab. Sparse high order FEM for elliptic sPDEs. *Comput. Methods Appl. Mech. Engrg.*, 198(13-14):1149–1170, 2009.

- [15] Boeing. Statistical summary of commercial jet airplane accidents worldwide operations 1959 - 2009. Technical report, Aviation Safety Boeing Commercial Airplanes, Seattle, Washington 98124-2207, U.S.A., July 2010.
- [16] S. Boucheron, G. Lugosi, and P. Massart. A sharp concentration inequality with applications. *Random Structures Algorithms*, 16(3):277–292, 2000.
- [17] P. Diaconis and D. Freedman. On the consistency of Bayes estimates. *Ann. Statist.*, 14(1):1–67, 1986. With a discussion and a rejoinder by the authors.
- [18] P. W. Diaconis and D. Freedman. Consistency of Bayes estimates for nonparametric regression: normal theory. *Bernoulli*, 4(4):411–444, 1998.
- [19] A. Doostan and H. Owhadi. A non-adapted sparse approximation of PDEs with stochastic inputs. 230(8):3015–3034, 2011.
- [20] R. F. Drenick. Model-free design of aseismic structures. *J. Eng. Mech. Div., Am. Soc. Civ. Eng.*, 96(4):483–493, 1970.
- [21] R. F. Drenick. Aseismic design by way of critical excitation. *J. Eng. Mech. Div., Am. Soc. Civ. Eng.*, 99(4):649–667, 1973.
- [22] R. F. Drenick. On a class of non-robust problems in stochastic dynamics. In B. L. Clarkson, editor, *Stochastic Problems in Dynamics*, pages 237–255. Pitman, 1977.
- [23] R. F. Drenick, P. C. Wang, C. B. Yun, and A. J. Philippacopoulos. Critical seismic response of nuclear reactors. *J. Nucl. Eng. Design*, 58(3):425–435, 1980.
- [24] E. B. Dynkin. Sufficient statistics and extreme points. *Ann. Prob.*, 6(5):705–730, 1978.
- [25] B. Efron and C. Morris. Stein’s paradox in statistics. *Scientific American*, 236(5):119–127, 1977.
- [26] I. H. Eldred, C. G. Webster, and P. G. Constantine. Design under uncertainty employing stochastic expansion methods. American Institute of Aeronautics and Astronautics Paper 2008–6001, 2008.
- [27] I. Elishakoff and M. Ohsaki. *Optimization and Anti-Optimization of Structures Under Uncertainty*. World Scientific, London, 2010.
- [28] L. Esteva. Seismic risk and seismic design. In *Seismic Design for Nuclear Power Plants*, pages 142–182. The M.I.T. Press, 1970.
- [29] R. Ghanem. Ingredients for a general purpose stochastic finite elements implementation. *Comput. Methods Appl. Mech. Engrg.*, 168(1-4):19–34, 1999.
- [30] R. Ghanem and S. Dham. Stochastic finite element analysis for multiphase flow in heterogeneous porous media. *Transp. Porous Media*, 32(3):239–262, 1998.

- [31] W. D. Gillford. Risk analysis and the acceptable probability of failure. *Structural Engineer*, 83(15):25–26, 2005.
- [32] H. J. Godwin. On generalizations of Tchebychef’s inequality. *J. Amer. Statist. Assoc.*, 50(271):923–945, 1955.
- [33] W. Hoeffding. On the distribution of the number of successes in independent trials. *Ann. Math. Statist.*, 27(3):713–721, 1956.
- [34] W. Hoeffding. The role of assumptions in statistical decisions. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. I*, pages 105–114, Berkeley and Los Angeles, 1956. University of California Press.
- [35] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58(301):13–30, 1963.
- [36] W. A. Hustrulid, M. McCarter, and D. J. A. Van Zyl, editors. *Slope Stability in Surface Mining*. Society for Mining Metallurgy & Exploration, 2001.
- [37] K. Isii. On a method for generalizations of Tchebycheff’s inequality. *Ann. Inst. Statist. Math. Tokyo*, 10(2):65–88, 1959.
- [38] K. Isii. The extrema of probability determined by generalized moments. I. Bounded random variables. *Ann. Inst. Statist. Math.*, 12(2):119–134; errata, 280, 1960.
- [39] K. Isii. On sharpness of Tchebycheff-type inequalities. *Ann. Inst. Statist. Math.*, 14(1):185–197, 1962/1963.
- [40] H. Joe. Majorization, randomness and dependence for multivariate distributions. *Ann. Probab.*, 15(3):1217–1225, 1987.
- [41] R. A. Johnson. Products of two Borel measures. *Trans. Amer. Math. Soc.*, 269(2):611–625, 1982.
- [42] S. Karlin and W. J. Studden. *Tchebycheff Systems: With Applications in Analysis and Statistics*. Pure and Applied Mathematics, Vol. XV. Interscience Publishers John Wiley & Sons, New York-London-Sydney, 1966.
- [43] A. F. Karr. Extreme points of certain sets of probability measures, with applications. *Math. Oper. Res.*, 8(1):74–85, 1983.
- [44] K. E. Kelly. The myth of  $10^{-6}$  as a definition of acceptable risk. In *Proceedings of the International Congress on the Health Effects of Hazardous Waste*, Atlanta, 1993. Agency for Toxic Substances and Disease Registry.
- [45] D. G. Kendall. Simplexes and vector lattices. *J. London Math. Soc.*, 37(1):365–371, 1962.

- [46] A.A. Kidane, A. Lashgari, B. Li, M. McKerns, M. Ortiz, H. Owhadi, G. Ravichandran, M. Stalzer, and T.J. Sullivan. Rigorous model-based uncertainty quantification with application to terminal ballistics. Part I: Systems with controllable inputs and small scatter. *Journal of the Mechanics and Physics of Solids*, 60(5):983–1001, 2012.
- [47] N. Lama, J. Wilsona, and G. Hutchinsona. Generation of synthetic earthquake accelograms using seismological modeling: a review. *Journal of Earthquake Engineering*, 4(3):321–354, 2000.
- [48] T. Leonard and J. S. J. Hsu. *Bayesian methods*, volume 5 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1999. An analysis for statisticians and interdisciplinary researchers.
- [49] J. Li and Z. An. Random function model research on strong ground motion. In *The 14th World Conference on Earthquake Engineering October 12-17, Beijing, China*, 2008.
- [50] J. S. Liu. *Monte Carlo strategies in scientific computing*. Springer Series in Statistics. Springer, New York, 2008.
- [51] A. L. Lopez, L. E. P. Rocha, D. L. Escobedo, and J. S. Sesma. Reliability and vulnerability analysis of electrical substations and transmission towers for definition of wind and seismic damage maps for Mexico. In *11th Americas Conference on Wind Engineering - San Juan, Puerto Rico*, 2009.
- [52] L. J. Lucas, H. Owhadi, and M. Ortiz. Rigorous verification, validation, uncertainty quantification and certification through concentration-of-measure inequalities. *Comput. Methods Appl. Mech. Engrg.*, 197(51-52):4591–4609, 2008.
- [53] N. Mantel and W. R. Bryan. “Safety” testing of carcinogenic agents. *J. Natl. Cancer Inst.*, 27:455–470, 1961.
- [54] A. W. Marshall and I. Olkin. Multivariate Chebyshev inequalities. *Ann. Math. Statist.*, 31(4):1001–1014, 1960.
- [55] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and its Applications*, volume 143 of *Mathematics in Science and Engineering*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1979.
- [56] S. Matsuda and K. Asano. Analytical expression of earthquake input energy spectrum based on spectral density model of ground acceleration. In *Proceeding of the 8th U. S. National Conference on Earthquake Engineering, April 18-22, San Francisco, USA*, 2006.
- [57] C. McDiarmid. On the method of bounded differences. In *Surveys in combinatorics, 1989 (Norwich, 1989)*, volume 141 of *London Math. Soc. Lecture Note Ser.*, pages 148–188. Cambridge Univ. Press, Cambridge, 1989.

- [58] C. McDiarmid. Concentration. In *Probabilistic methods for algorithmic discrete mathematics*, volume 16 of *Algorithms Combin.*, pages 195–248. Springer, Berlin, 1998.
- [59] M. McKerns, P. Hung, and M. Aivazis. *Mystic*: A simple model-independent inversion framework, 2009. <http://dev.danse.us/trac/mystic>.
- [60] M. McKerns, H. Owhadi, C. Scovel, T. J. Sullivan, and M. Ortiz. The optimal uncertainty algorithm in the *Mystic* framework. Technical report, Caltech, August 2011. CACR Technical Publication, <http://arxiv.org/pdf/1202.1055v1>.
- [61] M. M. McKerns, L. Strand, T. J. Sullivan, A. Fang, and M. A. G. Aivazis. Building a framework for predictive science. In *Proceedings of the 10th Python in Science Conference (SciPy 2011)*, 2011. <http://arxiv.org/pdf/1202.1056v1>.
- [62] H. P. Mulholland and C. A. Rogers. Representation theorems for distribution functions. *Proc. London Math. Soc. (3)*, 8(2):177–223, 1958.
- [63] N. M. Newmark and E. Rosenblueth. *Fundamentals of Earthquake Engineering*. Prentice Hall, 1972.
- [64] H. Niederreiter. *Random number generation and quasi-Monte Carlo methods*, volume 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
- [65] I. Olkin and J. W. Pratt. A multivariate Tchebycheff inequality. *Ann. Math. Statist.*, 29(1):226–234, 1958.
- [66] H. Owhadi. Anomalous slow diffusion from perpetual homogenization. *Ann. Probab.*, 31(4):1935–1969, 2003. <http://dx.doi.org/10.1214/aop/1068646372>.
- [67] P. M. Pardalos and S. A. Vavasis. Quadratic programming with one negative eigenvalue is NP-hard. *J. Global Optim.*, 1(1):15–22, 1991.
- [68] R. R. Phelps. *Lectures on Choquet’s Theorem*, volume 1757 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, second edition, 2001.
- [69] I. Pinelis. Exact inequalities for sums of asymmetric random variables, with applications. *Probab. Theory Related Fields*, 139(3-4):605–635, 2007.
- [70] I. Pinelis. On inequalities for sums of bounded random variables. *J. Math. Inequal.*, 2(1):1–7, 2008.
- [71] M. J. D. Powell. A direct search optimization method that models the objective and constraint functions by linear interpolation. In S. Gomez and J. P. Hennart, editors, *Advances in Optimization and Numerical Analysis*, page 5167. Kluwer Academic, 1994.



- [72] K. V. Price, R. M. Storn, and J. A. Lampinen. *Differential Evolution: A Practical Approach to Global Optimization*. Natural Computing Series. Springer-Verlag, Berlin, 2005.
- [73] A. D. Rikun. A convex envelope formula for multilinear functions. *J. Global Optim.*, 10(4):425–437, 1997.
- [74] R. T. Rockafellar. Augmented Lagrange multiplier functions and duality in non-convex programming. *SIAM J. Control*, 12(2):268–285, 1974.
- [75] A. Saltelli, K. Chan, and E. M. Scott, editors. *Sensitivity Analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester, 2000.
- [76] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global Sensitivity Analysis. The Primer*. John Wiley & Sons Ltd., Chichester, 2008.
- [77] S. M. Samuels. On a Chebyshev-type inequality for sums of independent random variables. *Ann. Math. Statist.*, 37(1):248–259, 1966.
- [78] C. Scovel and I. Steinwart. Hypothesis testing for validation and certification. *Journal of Complexity*, Submitted, 2010.
- [79] X. Shen and L. Wasserman. Rates of convergence of posterior distributions. *Ann. Statist.*, 29(3):687–714, 2001.
- [80] H. D. Sherali. Convex envelopes of multilinear functions over a unit hypercube and over special discrete sets. *Acta Math. Vietnam.*, 22(1):245–270, 1997.
- [81] I. H. Sloan. Sparse sampling techniques. Presented at 2010 ICMS Uncertainty Quantification workshop, 2010.
- [82] I. H. Sloan and S. Joe. *Lattice methods for multiple integration*. Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1994.
- [83] H. M. Soekkhla, editor. *Aviation Safety: Human Factors, System Engineering, Flight Operations, Economics, Strategies, Management*. VSP, 1997.
- [84] MA Software. Matruss 1.3, 2008. <http://www.soft3k.com/MATruss-p4424.htm>.
- [85] A. Srivastav and P. Stangier. Algorithmic Chernoff–Hoeffding inequalities in integer programming. *Random Structures Algorithms*, 8(1):27–58, 1996.
- [86] P. J. Stafford, J. B. Berrill, and J. R. Pettinga. New predictive equations for arias intensity from crustal earthquakes in new zealand. *Journal of Seismology*, 13(1):31–52, 2009.
- [87] S. Stein and M. Wyssession. *An Introduction to Seismology, Earthquakes, and Earth Structure*. Wiley-Blackwell, 2002.

- [88] R. M. Storn and K. V. Price. Differential evolution — a simple and efficient heuristic for global optimization over continuous spaces. *J. Global Optim.*, 11(4):341–359, 1997.
- [89] A. M. Stuart. Inverse problems: a Bayesian perspective. *Acta Numer.*, 19:451–559, 2010.
- [90] T. J. Sullivan, M. McKerns, D. Meyer, F. Theil, H. Owhadi, and M. Ortiz. Optimal uncertainty quantification for legacy data observations of lipschitz functions. *arXiv:1202.1928*, 2012.
- [91] T. J. Sullivan, U. Topcu, M. McKerns, and H. Owhadi. Uncertainty quantification via codimension-one partitioning. *Int. J. Num. Meth. Engrg.*, 85(12):1499–1521, 2011.
- [92] I. Takewaki. Seismic critical excitation method for robust design: A review. *J. Struct. Engrg.*, 128(5):665–672, 2002.
- [93] I. Takewaki. *Critical Excitation Methods in Earthquake Engineering*. Elsevier Science, 2007.
- [94] R. A. Todor and C. Schwab. Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients. *IMA J. Numer. Anal.*, 27(2):232–261, 2007.
- [95] F. Topsøe. *Topology and Measure*. Lecture Notes in Mathematics, Vol. 133. Springer-Verlag, Berlin, 1970.
- [96] H. Trainsson, A. S. Kiremidjian, and S. Winterstein. Modeling of earthquake ground motion in the frequency domain. Technical Report 134, Department of Civil and Environmental Engineering, Stanford University, 2000.
- [97] J. W. van de Lindt and G.-H. Goh. Earthquake duration effect on structural reliability. *J. Struct. Engrg.*, 130(5):821–826, 2004.
- [98] L. Vandenberghe, S. Boyd, and K. Comanor. Generalized Chebyshev bounds via semidefinite programming. *SIAM Rev.*, 49(1):52–64 (electronic), 2007.
- [99] H. von Weizsäcker and G. Winkler. Integral representation in the set of solutions of a generalized moment problem. *Math. Ann.*, 246(1):23–32, 1979/80.
- [100] G. Winkler. Extreme points of moment sets. *Math. Oper. Res.*, 13(4):581–587, 1988.
- [101] D. Xiu. Fast numerical methods for stochastic computations: a review. *Commun. Comput. Phys.*, 5(2-4):242–272, 2009.